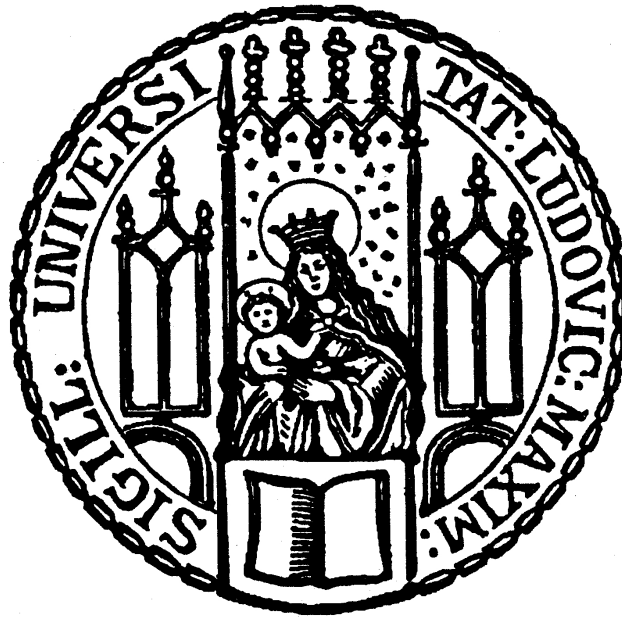


Study of the $Z + b\bar{b}$ Production with the ATLAS Detector at the LHC

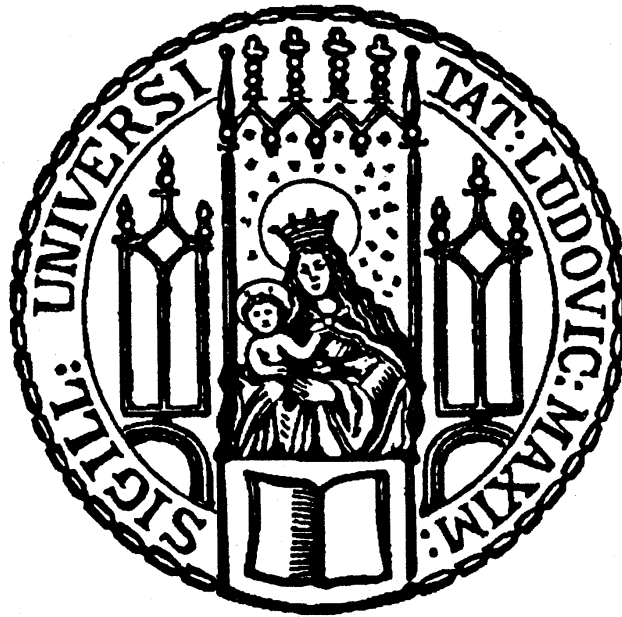


Dissertation der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Dan Nicolae Vlădoiu
geboren in Bukarest

München, Juli 2013

Study of the $Z + b\bar{b}$ Production with the ATLAS Detector at the LHC



Dissertation der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Dan Nicolae Vlădoiu
geboren in Bukarest

München, Juli 2013

Erstgutachterin: Prof. Dr. Dorothee Schaile
Zweitgutachter: Prof. Dr. Martin Faessler

Tag der mündlichen Prüfung: 01. Oktober 2013

Abstract

A first inclusive measurement of the production cross-section for a Z boson in association with a bottom-antibottom quark pair at $\sqrt{s} = 7$ TeV at the Large Hadron Collider is presented, using the full dataset of $\int L(t) dt = 4.64 \text{ pb}^{-1}$ collected by the ATLAS detector during the year 2011.

Such a measurement provides a direct test of perturbative quantum chromodynamics predictions, and, for production of a Higgs boson in association with a Z boson, $Z + b\bar{b}$ represents the highest irreducible background.

The decay channels of the Z boson considered are: $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$. For the reconstruction of the Z boson candidate, exactly two leptons with opposite charge and of the same-flavor are required to have an invariant mass within 76 GeV - 106 GeV. Jets with weakly decaying b -hadrons are selected with a b -tagging algorithm at 75% efficiency. For the reduction of the top-antitop background, events are required to have a missing transverse energy of less than 30 GeV.

The extraction of the signal cross-section from the data sample was realized with a binned likelihood template fit method, using as template fit a strong discriminant between heavy-flavor and charm/light-flavor jets. Three exclusive templates are produced: “ bb ” – events with a Z boson candidate and at least two identified b -jets, obtained from simulated samples with a Z boson and heavy b -jets; “ $non - bb$ ” – events with a Z boson candidate and less than two identified b -jets, derived from simulated samples with a Z boson and light-jets, and “ $other$ ” - events with a Z boson candidate and any number of b -jets created from simulated top-antitop and diboson samples. The sum of the three template distributions was fit to the data keeping the normalization of the $other$ template fixed and floating the normalization of the bb and $non - bb$ templates.

The measured cross-section results in the electron and muon channels are:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow e^+e^-) = 436.3 \pm 32.4 \text{ (stat.) } {}^{+52.3}_{-50.3} \text{ (syst.) } \pm 7.9 \text{ (lumi.) fb,}$$

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow \mu^+\mu^-) = 552.0 \pm 33.6 \text{ (stat.) } {}^{+64.4}_{-59.6} \text{ (syst.) } \pm 10.0 \text{ (lumi.) fb.}$$

For the combination of the two channels, the electron and muon channel templates are added together and fit to the sum of the data. The cross-section for the combined channel is:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 499.1 \pm 23.5 \text{ (stat.) } {}^{+57.7}_{-53.2} \text{ (syst.) } \pm 8.9 \text{ (lumi.) fb,}$$

in reasonable agreement with a next-to-leading order theoretical prediction:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 403 {}^{+82}_{-66} \text{ fb.}$$

Zusammenfassung

Eine inklusive Messung des Produktionswirkungsquerschnitts eines Z-Bosons in Assoziation mit einem Bottom-Antibottom Quarkpaar bei $\sqrt{s} = 7\text{ TeV}$ am Large Hadron Collider wird vorgestellt. Hierfür wird der gesamte im Jahr 2011 vom ATLAS-Detektor aufgenommene Datensatz von $\int L(t) dt = 4.64\text{ pb}^{-1}$ verwendet.

Eine derartige Messung ist interessant für einige Bereiche der Hochenergiephysik. Man erhält beispielsweise einen direkten Test von Vorhersagen der perturbativen Quantenchromodynamik. Zudem stellt $Z + b\bar{b}$ für die Produktion eines Higgs-Bosons in Assoziation mit einem Z-Boson den grössten irreduziblen Untergrund dar.

Die hier betrachteten Zerfallskanäle des Z-Bosons sind $Z \rightarrow e^+e^-$ und $Z \rightarrow \mu^+\mu^-$. Für die Rekonstruktion des Z-Bosonkandidaten werden genau zwei Leptonen mit gleichem Flavour und entgegengesetzter elektrischer Ladung mit einer invarianten Masse zwischen 76 GeV und 106 GeV gefordert. Jets aus schwach zerfallenden b-Hadronen werden mit einem b-tagging Algorithmus bei 75% Effizienz ausgewählt. Zur Reduzierung des Top-Antitop-Untergrunds sollen die ausgewählten Ereignisse weniger als 30 GeV fehlende Transversalenergie aufweisen.

Der Signalwirkungsquerschnitt wird aus den Daten mithilfe einer “binned likelihood template fit method” gewonnen. Der Input basiert in jedem Kanal auf der Simulation, als Fit-Variable wird eine starke Diskriminante zwischen Jets aus schweren Quarks und Jets aus Charm-Quarks oder leichten Quarks verwendet.

Drei exklusive Templates werden betrachtet: “*bb*” Ereignisse mit einem Z-Bosonkandidaten und mindestens zwei identifizierten b-Jets; “*nonbb*” Ereignisse mit einem Z-Bosonkandidaten und weniger als zwei identifizierten b-Jets, jeweils gewonnen aus Datensätzen mit einem Z-Boson und schweren bzw. leichten Jets; “*other*” Ereignisse mit einem Z-Bosonkandidaten und einer beliebigen Anzahl an b-Jets, gewonnen aus Top-Antitop-Datensätzen und Diboson-Datensätzen. Die Summe der drei Verteilungen wird bei fester Normalisierung des *other* Templates und variabler Normalisierung der *bb* und *non - bb* Templates an die Daten angepasst.

Die gemessenen Wirkungsquerschnitte im Elektron- und Myonkanal sind:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow e^+e^-) = 436.3 \pm 32.4 \text{ (stat.) } {}^{+52.3}_{-50.3} \text{ (syst.) } \pm 7.9 \text{ (lumi.) fb,}$$

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow \mu^+\mu^-) = 552.0 \pm 33.6 \text{ (stat.) } {}^{+64.4}_{-59.6} \text{ (syst.) } \pm 10.0 \text{ (lumi.) fb.}$$

Für die Kombination der beiden Kanäle werden die Templates des Elektron- und des Myonkanals addiert und an die Summe der Daten angepasst. Für den Wirkungsquerschnitt des kombinierten Kanals ergibt sich:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 499.1 \pm 23.5 \text{ (stat.) } {}^{+57.7}_{-53.2} \text{ (syst.) } \pm 8.9 \text{ (lumi.) fb,}$$

in Einklang mit der theoretischen Vorhersage in nächstführender Ordnung von:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+ l^-) = 403_{-66}^{+82} \text{ fb.}$$

Für meinen ersten Physiklehrer, meinen Vater.

Contents

1	Short History of Particle Physics	15
2	Introduction	18
2.1	The Standard Model of Particle Physics	18
2.1.1	The $SU(2)_L \otimes U(1)_Y$ Formalism of Electroweak Interactions . .	20
2.1.2	The $SU(3)_C$ Formalism of Strong Interactions	26
2.1.3	Production of $Z + b\bar{b}$ events	30
2.1.4	Production of ZH events	30
2.2	Experimental Ingredients	33
2.2.1	Luminosity and Cross-Section Measurements	33
2.2.2	Parton Distribution Functions	34
2.2.3	Pileup, Multiparton Interactions and the Underlying Event	35
2.2.4	Parton Shower and Hadronization Models	37
3	Experimental Setup	38
3.1	The Large Hadron Collider	38
3.2	The ATLAS Detector	40
3.3	The Inner Detector	41
3.4	The Solenoid Magnet	44
3.5	The Calorimeter System	45
3.6	The Toroid Magnet	47
3.7	The Muon Spectrometer	48
3.8	Trigger and Data Acquisition	51
3.9	ATLAS Forward Detectors	52
3.10	The Worldwide LHC Computing Grid	53
3.11	MonteCarlo Simulation	54

4	Data and MonteCarlo samples	56
4.1	Data Samples	56
4.2	MonteCarlo Samples	57
5	Object Reconstruction and Event Selection	63
5.1	Reconstruction and Event Preselection	63
5.2	Final Event Selection	72
6	Multijet Background Estimation	91
6.1	Electron Channel	95
6.2	Muon Channel	100
6.3	Multijets Systematic Uncertainty	106
7	Cross-Section Measurement	107
7.1	Description of the Template Fit	107
7.2	Selection Efficiencies Estimation	111
7.3	Evaluation of the Systematic Uncertainties	114
7.4	Results	120
7.5	Theoretical Predictions	123
8	Summary and Conclusions	125
A	Appendix	127
B	Appendix	131

1 Short History of Particle Physics

The intuitive idea according to which the physical universe surrounding us has a discrete intrinsic nature goes back to the antiquity. It is formally accepted nowadays that the atom concept (from greek *ἄτομος* , *uncuttable*) was born in the Greek city of Abdera, Thrace, and was promoted by the greek philosopher Leucippus and his followers such as Democritus, Epicurus and Lucretius. Even though there are fundamental differences between the modern concept of the atom and the ancient one, some remnants of the philosophical model survived time such as, for instance, the fact that objects are composed of a vast number of atoms, separated by empty space, which come in relative small number of types.

One of the first scientific models of the atom was proposed in the year 1661 by the British natural philosopher Robert Boyle who argued for the existence of atoms and atom-clusters (called later *molecules* by Amadeo Avogadro) which are to be found in a continuous motion and provide the building blocks of matter. Later on, in 1789, Antoine Lavoisier proposed the *law of conservation of mass* and sustained that matter is composed of *elements* which have a discrete substructure and cannot be further reduced via chemical methods. Two decades later, John Dalton elaborated the *law of multiple proportions* and argued that every element is made up of a unique type of atom. In 1865 the Austrian scientist Johann Josef Loschmidt set an important milestone by estimating for the first time the size of the air molecules.

During his “annus mirabilis”, 1905, Albert Einstein, based on Planck’s hypothesis of electromagnetic field quanta (1901), explains the empirically puzzling result of the *photoelectric* effect. In the same year, Einstein, grounding his arguments on the kinetic theory of gasses, shed light upon yet another empirical riddle, the *brownian motion*.

The twentieth century is the time of great discoveries, unprecedented in the history of science, which had fundamental impact on our understanding of the Universe at the very large (General Relativity) and very small (Quantum Mechanics) scales.

The end of the nineteenth century was already announcing what was yet to come. First the electron (from greek *ἤλεκτρον*, *amber*, term proposed in 1891 by George Johnstone Stoney) was discovered in the year 1896 by J.J. Thompson and his colleagues John S. Townsend and H. A. Wilson by deflecting cathodic rays inside a Crooks tube via electric and magnetic fields. Thirteen years later, Millikan in his famous “oil-drop experiment”

measured the electric charge, and consequently the electron's mass.

By bombarding high-purity gaseous nitrogen with alpha particles in 1917, Rutherford discovered the proton (from greek $\pi\rho\acute{o}\tau\omicron\nu$, *the first*) via what became to be known today as the first nuclear reaction: ${}^4_2\alpha + {}^{14}_7\text{N} \rightarrow {}^{18}_9\text{F}^* \rightarrow {}^{17}_8\text{O} + {}^1_1\text{p}$. The first antiparticle, the positron, is discovered in 1931 by Andersen. One year later Sir James Chadwick discovered the neutron (from the latin root *neutral* plus the greek ending *on*) in the experimental physics laboratory of Cambridge University by bombarding targets of nitrogen, oxygen, helium and argon with alpha particles. The kinematical analysis of the recoils suggested that there had to be a neutral massive particle and not gamma rays as was initially assumed. In the same year Heisenberg introduces the formalism of isospin, treating the proton and the neutron as two projections of a unique particle - the nucleon, using for the first time the $SU(2)$ group to particle physics.

In the next three decades that followed, the group of “elementary” particles increases tremendously by the discovery of the neutrino, the muon, tauon, and a vast number of mesons and baryons. It is said that Carl David Anderson jokingly exclaimed once: “the finder of a new elementary particle used to be rewarded by a Nobel Prize, but such a discovery now ought to be punished by a 10000 dollar fine.”

A significant simplification came in 1964 with the *quark model* formalism (the term quark was inspired to Gell-Mann by James Joyce's novel “Finnegans Wake”), grounded on the $SU(3)_C$ group, proposed independently by Murray Gell-Mann and George Zweig. According to the quark model hadrons are not elementary particles per se, but compound systems of *quarks* and *gluons*.

The theory unifying the electromagnetic and weak interactions (responsible for radioactive beta decays) was finalized around 1967, by Sheldon Glashow, Abdus Salam and Steven Weinberg. The model, based on the $SU(2)_L \otimes U(1)_Y$ symmetry group, predicted the existence of three additional massive vector bosons, besides the massless photon, namely the W^+ , W^- and Z^0 . Their first direct observation came in 1983, at CERN's SPS (Super Proton Synchrotron, a proton-antiproton collider), by the UA1 and UA2 experiments.

Since the gauge structure of the $SU(2)_L \otimes U(1)_Y$ group requires massless vector bosons, the symmetry has to be broken by an additional unknown mechanism. Three famous papers ([1, 2, 3]) published in the period 1963-1964 by F. Englert, R. Brout and P. Higgs provided a possible way of breaking the $SU(2)_L \otimes U(1)_Y$ symmetry and giving mass to the vector bosons via what is nowadays colloquially called the *Higgs mechanism*. A considerable amount of theoretical and especially experimental effort will be invested in the search of the elusive scalar boson during the following five decades. In the course of the year 2012 both the ATLAS and CMS experiments at the LHC, reported on the discovery of a new particle compatible with the Standard Model (SM) scalar boson ([4, 5]). Subsequent searches, using an increased amount of integrated luminosity have brought additional evidence for the SM Higgs boson scenario [6, 7].

Whether the signal observed by the two experiments is the long searched SM scalar boson required for the electroweak symmetry breaking or not, this is yet to be revealed in future

by detailed measurements of the production and decay channels and comparison with the SM Higgs boson predictions. A first step will be a good understanding of the background channels out of which $Z + b\bar{b}$ production, the topic of this thesis, plays an important role. Besides the possible observation of the SM Higgs, the ATLAS and CMS experiments were designed for a much broader domain of searches such as supersymmetry, extra dimensions, evidence of new CP-violation mechanisms or even new massive vector bosons such as the Z' or W' .

Counting the Higgs boson, the Standard Model as it presents itself today, gives a total of 61 elementary particles. If the new Tera electron-volt epoch initiated by the Tevatron and LHC will bring a simplification to the concept of “elementarity” in particle physics, or, additional fundamental constituents, is a question that still waits for an answer.

2 Introduction

2.1 The Standard Model of Particle Physics

The Standard Model unifies the electromagnetic and weak processes into a single electroweak framework, and, together with the strong force, describes with high precision the dynamics of the known elementary particles via the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge group.

Taking as a first criterion of classification the spin, the elementary particles of the Standard Model can be divided into gauge bosons, spin-1 field quanta responsible for transmitting the gauge interactions, and fermions, spin-1/2 particles, subjects of the interaction itself. A systemization of the elementary fermions can be realized further by joining together the weak isospin partners of the six known quarks and leptons into three families (see tables 2.1, 2.2 and 2.3):

Flavor	Electric Charge [e]	Mass [MeV]	Spin	Interaction
ν_e	0	$\neq 0$	1/2	weak
e	-1	0.511	1/2	electroweak
u	+2/3	≈ 2.3	1/2	strong, electroweak
d	-1/3	≈ 4.8	1/2	strong, electroweak

Table 2.1: First generation of quarks and leptons [8].

Flavor	Electric Charge [e]	Mass [MeV]	Spin	Interaction
ν_μ	0	$\neq 0$	1/2	weak
μ	-1	≈ 105	1/2	electroweak
c	+2/3	$\approx 1,275$	1/2	strong, electroweak
s	-1/3	≈ 95	1/2	strong, electroweak

Table 2.2: Second generation of quarks and leptons [8].

Flavor	Electric Charge [e]	Mass [GeV]	Spin	Interaction
ν_τ	0	$\neq 0$	1/2	weak
τ	-1	≈ 1.778	1/2	electroweak
t	+2/3	≈ 173.5	1/2	strong, electroweak
b	-1/3	≈ 4.18	1/2	strong, electroweak

Table 2.3: Third generation of quarks and leptons [8].

A classification of the Standard Model gauge bosons can be realized via the type of interaction transmitted as shown in table 2.4.

Symbol	Spin	Mass [MeV]	Full Width [MeV]	Interaction
g	1	null	0	strong
Z^0	1	≈ 90.2 GeV	2495	weak
W^\pm	1	≈ 80.4 GeV	2085	electroweak
γ	1	null	0	electromagnetic

Table 2.4: The gauge bosons of the Standard Model [8].

The Standard Model as a relativistic quantum field theory, has been subject to a continuous process of development and improvement. The mathematical structure of the model has been established via trial and error over a period of approximately two decades, and cluttered in a form close to its present form in 1967 when S. Weinberg and A. Salam incorporated the Higgs mechanism to Glashow's version of the electroweak theory (the model is sometimes referred to as the GWS theory). After providing a way to hide and break the electroweak symmetry, the Standard Model proved to be a remarkable accurate theory, exhibiting two of the most important ingredients of a good empirical theory: accordance with existing experimental data and new verifiable predictions. The discovery of the weak neutral current in 1973 at CERN (in $\bar{\nu}_\mu e$ elastic scattering), those of the massive gauge bosons (W^+ , W^- and Z^0) in 1983, the discovery of the charm quark and later

those of bottom and top quarks (a third generation of quarks was required after the discovery of CP-violation in kaons), made the Standard Model the most accurate description of elementary particle physics.

In the next two sub-sections a brief description of the $SU(2)_L \otimes U(1)_Y$ electroweak theory is presented followed by the $SU(3)_C$ formalism of strong interactions. The theoretical framework is based on [9, 10]. Finally the last two sub-sections are dedicated to the $Z + b\bar{b}$ and ZH production at the Large Hadron Collider.

2.1.1 The $SU(2)_L \otimes U(1)_Y$ Formalism of Electroweak Interactions

One of the first attempts to describe weak interactions was proposed by Enrico Fermi as a theoretical framework of the β^\pm decay of radioactive nuclei. According to Fermi's model, weak processes take place via a point interaction and require no exchanged particle. The matrix element of such an interaction can be written as:

$$\mathfrak{M}^{\text{ch}} = G_F \cdot (\bar{u}_2 \gamma^\mu u_1) \cdot (\bar{u}_4 \gamma_\mu u_3), \quad (2.1)$$

where G_F represents Fermi's constant and has an approximate value of:

$$G_F \approx 1,16637 \cdot 10^{-5} \cdot \text{GeV}^{-2}. \quad (2.2)$$

This is different from QED (quantum electrodynamics) processes, where the interaction occurs via the exchange of virtual photons. In the lowest order, the interaction is approximated via the exchange of a single photon:

$$\mathfrak{M}^{\text{em}} \propto (j^\mu)^{\text{em}} \cdot \left(\frac{-g_{\mu\nu}}{q^2} \right) \cdot (j^\nu)^{\text{em}}, \quad \text{where: } (j^\sigma)^{\text{em}} = \bar{u} \cdot \gamma^\sigma \cdot u. \quad (2.3)$$

where $g_{\mu\nu}$ is the metric tensor and q corresponds to the 4-momentum carried by the virtual photon.

Equation (2.1) describes a 4-particle point interaction, where u_1 , \bar{u}_2 , u_3 , and \bar{u}_4 are Dirac spinors:

$$u^\uparrow = \lambda \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x + i p_y}{E+m} \end{pmatrix} \quad u^\downarrow = \lambda \begin{pmatrix} 0 \\ 1 \\ \frac{p_x - i p_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix} \quad v^\uparrow = \lambda \begin{pmatrix} \frac{p_x - i p_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} \quad v^\downarrow = -\lambda \begin{pmatrix} \frac{p_z}{E+m} \\ \frac{p_x + i p_y}{E+m} \\ 1 \\ 0 \end{pmatrix} \quad (2.4)$$

with $\lambda = \sqrt{E+m}$ and $\bar{u}_i = u_i^\dagger \cdot \gamma^0$. The spinors describe particles with spin up u^\uparrow or down u^\downarrow and antiparticles with spin up v^\uparrow and down v^\downarrow respectively.

The structure of the amplitude \mathfrak{M}^{ch} resembles the known Lorentz invariant electromagnetic amplitude \mathfrak{M}^{em} , with the above mentioned caveat.

The limits of this model became obvious by the discovery of parity violation, theorized by T. D. Lee and C. N. Yang in 1956 and confirmed experimentally by C.S. Wu one year later. This required a modification to Fermi's original model by replacing the vectorial term with a vector - axial term:

$$\mathfrak{M}^{\text{ch}} = \frac{G_F}{\sqrt{2}} \cdot [\bar{u}_2 \gamma^\mu (1 - \gamma^5) u_1] \cdot [\bar{u}_4 \gamma_\mu (1 - \gamma^5) u_3]. \quad (2.5)$$

The $\gamma^\mu (1 - \gamma^5)$ factor will bring in an additional change in the nature of weak charged interactions restricting their action only to left-handed chiral fermions or right-handed chiral antifermions. Finally, in accordance with the GWS model and incorporating massive gauge bosons, the interaction took the form:

$$\mathfrak{M}^{\text{ch}} = \frac{g_w^2}{8} \cdot [\bar{u}_2 \gamma^\mu (1 - \gamma^5) u_1] \cdot \frac{g_{\mu\nu}}{M_W^2 - q^2} \cdot [\bar{u}_4 \gamma^\nu (1 - \gamma^5) u_3]. \quad (2.6)$$

In the modern formalism, G_F is replaced by the dimensionless coupling constant g_w and the process takes place via a charged weak current mediated by a non-zero mass vector boson. In the low energy limit, $q^2 \ll M_W^2$, Fermi's model is recovered.

Neutral weak processes take place via a similar amplitude:

$$\mathfrak{M}^{\text{n}} = \frac{g_z^2}{4} \cdot [\bar{u}_2 \gamma^\mu (c_V^i - c_A^i \gamma^5) u_1] \cdot \frac{g_{\mu\nu}}{M_Z^2 - q^2} \cdot [\bar{u}_4 \gamma^\nu (c_V^j - c_A^j \gamma^5) u_3], \quad (2.7)$$

where $c_V^{i,j}$ and $c_A^{i,j}$ are the vectorial and axial coefficients specific to a given lepton or quark flavor.

According to the $SU(2)_L \otimes U(1)_Y$ electroweak formalism, the massive weak gauge bosons, W^+ , W^- and Z^0 share a common origin with the massless photon. The model assumes the existence of a weak isovector field W_μ^i , with coupling constant g and of a weak isoscalar field B_μ , with coupling constant g' , where both sets of fields are required to be massless under the virtue of local gauge invariance. The associated currents for the isovector components are defined as:

$$j_\mu^i = \bar{\chi}_L \gamma_\mu \frac{\tau_i}{2} \chi_L, \quad \text{with: } i = 1, 2, 3; \quad (2.8)$$

where:

$$\chi_L^l = \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L \quad \text{and:} \quad \chi_L^q = \begin{pmatrix} u \\ d \end{pmatrix}_L. \quad (2.9)$$

χ_L^l and χ_L^q describe the lepton and quark left-handed weak isospin doublet. The subscript "L" emphasises the left-handed nature of the weak interaction and the τ_i terms symbolise

Pauli's matrices. Since weak interactions only act upon left handed-fermions, the right-handed states remain invariant and form weak isoscalars such as l_R for leptons and u_R or d_R for quarks.

The corresponding current associated with the weak isoscalar field, A_μ , forms together with j_μ^3 two orthogonal states, provided it takes the form:

$$j_\mu^Y = 2(j_\mu^{em} - j_\mu^3), \quad (2.10)$$

replicating the Gell-Mann Nishijima relation from hadron multiplets systemization:

$$Q = I_3 + \frac{Y}{2}. \quad (2.11)$$

where Q is the electric charge, I_3 and Y are the third component of the isospin and the hypercharge respectively.

The general electroweak interaction term will be expressed by:

$$-i [g \cdot j_\mu^i (W^\mu)^i + \frac{g'}{2} \cdot j_\mu^Y A^\mu]. \quad (2.12)$$

In order to break the $SU(2)_L \otimes U(1)_Y$ symmetry and give mass to the weak vector bosons (and possibly to fermions) a field with non-zero vacuum expectation value, $v \neq 0$, was introduced ad-hoc. This field takes the form of a two-component complex scalar field with weak hypercharge +1:

$$h = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} h_1 + ih_2 \\ h_3 + ih_4 \end{pmatrix}, \quad (2.13)$$

and is described by the following Lagrangian:

$$\mathcal{L} = (\partial^\mu h)^* (\partial_\mu h) - V(h^* h), \quad (2.14)$$

Requiring local invariance under $SU(2)_L \otimes U(1)_Y$ transformations for the scalar field:

$$h \rightarrow e^{i(\theta_{(x)} \cdot \frac{\tau^i}{2} + \phi_{(x)} \cdot Y)} h, \quad (2.15)$$

the Lagrangian becomes:

$$\mathcal{L} = (\mathcal{D}^\mu h)^* (\mathcal{D}_\mu h) - V(h^* h), \quad (2.16)$$

where:

$$\mathcal{D}_\mu = \partial_\mu + i \frac{g'}{2} A_\mu Y + i \frac{g}{2} \tau^i \cdot W_\mu^i, \text{ with } i = 1, 2, 3. \quad (2.17)$$

is the covariant derivative. The scalar field potential is chosen as (see Figure 2.1):

$$V(h^*h) = \mu^2(h^*h) + |\lambda|(h^*h)^2. \quad (2.18)$$

By forcing the μ^2 parameter to be negative, the Lagrangian will automatically exhibit spontaneous symmetry breaking. Excitations of the scalar field, $h = \frac{1}{\sqrt{2}} \begin{pmatrix} h^+ \\ h^0 \end{pmatrix}$ (where $h^+ = h_1 + ih_2$ and $h^0 = h_3 + ih_4$), can be expanded as fluctuations about the ground state:

$$h = \begin{pmatrix} 0 \\ (v + \epsilon)/\sqrt{2} \end{pmatrix}, \quad (2.19)$$

where:

$$h_{ground} = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} \text{ and: } v = \sqrt{-\mu^2/|\lambda|}, \quad (2.20)$$

The h_1 and h_2 components have been chosen to be null in order to preserve the neutrality of vacuum as well as to eliminate non-physical oscillation terms between the field's components, while the h_4 component can be removed by a gauge transformation. This particular vacuum state will break the $SU(2)_L \otimes U(1)_Y$ gauge symmetry to the electromagnetic $U(1)_{em}$ gauge symmetry, meaning that all four initial generators (three for the isovector and one for the isoscalar) will no longer preserve their invariance after spontaneous symmetry breaking occurred. However, the electric charge remains an unbroken generator and consequently the photon emerges as a massless linear combination of A_μ and W_3 fields:

$$\gamma = B \cos \theta_w + W_3 \sin \theta_w, \text{ with mass: } M_\gamma = 0, \quad (2.21)$$

where θ_w is the weak mixing angle defined as: $\tan \theta_w = g'/g$. The Z^0 comes out as a massive combination of the same fundamental fields:

$$Z^0 = -B \sin \theta_w + W_3 \cos \theta_w, \text{ with: } M_{Z^0} = \frac{g \cdot v}{2 \cos \theta_w}, \quad (2.22)$$

Finally, the linear combination of the first two components of the weak isovector mix into the charged W^\pm bosons:

$$W^+ = \frac{1}{\sqrt{2}} (W_1 - i \cdot W_2), \quad W^- = \frac{1}{\sqrt{2}} (W_1 + i \cdot W_2) \text{ with mass: } M_{W^\pm} = \frac{g \cdot v}{2}. \quad (2.23)$$

Before the spontaneous symmetry breaking, the scalar field had a total of four degrees of freedom; three of them are incorporated now as longitudinal components of the massive gauge bosons. The fourth, acquiring mass via the same process, is identified with the scalar boson having a free (unpredicted by theory) mass parameter:

$$M_h = \sqrt{-2\mu^2}. \quad (2.24)$$

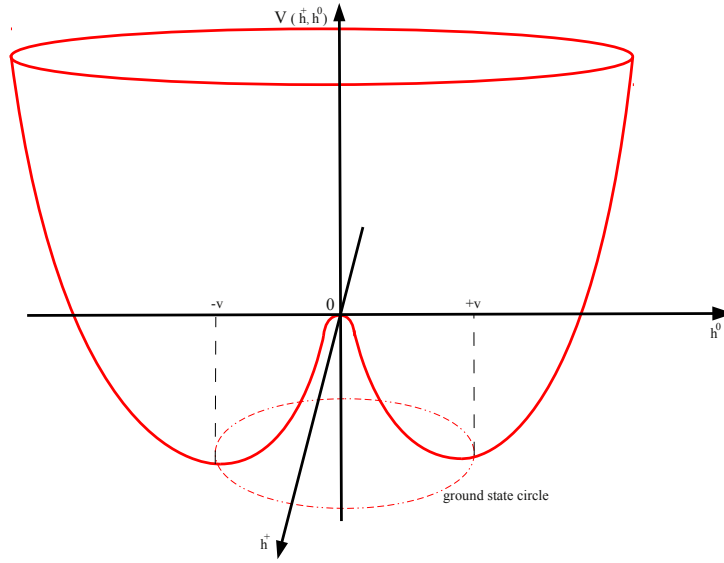


Figure 2.1: The spontaneous symmetry breaking potential of the scalar field.

The fundamental couplings can be assessed by adding free Dirac Lagrangian terms to the scalar Lagrangian from relation (2.14) and requiring local $SU(2)_L \otimes U(1)_Y$ gauge invariance to the combined Lagrangian. Figure 2.2 shows the fundamental coupling terms between the gauge bosons and fermions while Figure 2.3 shows the coupling terms between the gauge bosons themselves. In Figure 2.4 the couplings between the scalar field and fermions¹, massive gauge bosons and selfcouplings of the scalar field are shown.

¹There is a way in which fermions can acquire mass in the Standard Model by the ad-hoc insertion in the Lagrangian of Yukawa terms such as: $\alpha_Y \bar{\psi} \psi$. However the fermion masses are not predictable by theory since the value of the coupling constants, α_Y , still remain unknown.

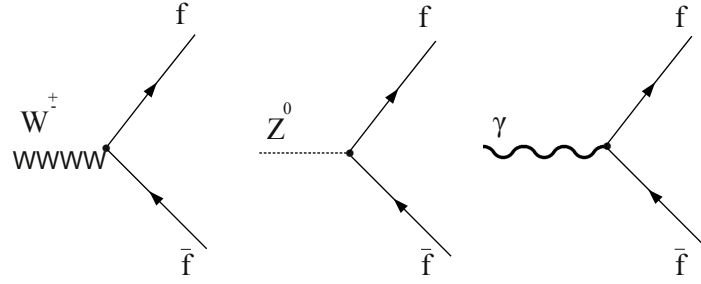


Figure 2.2: Coupling terms between $SU(2)_L \otimes U(1)_Y$ gauge bosons and fermions.

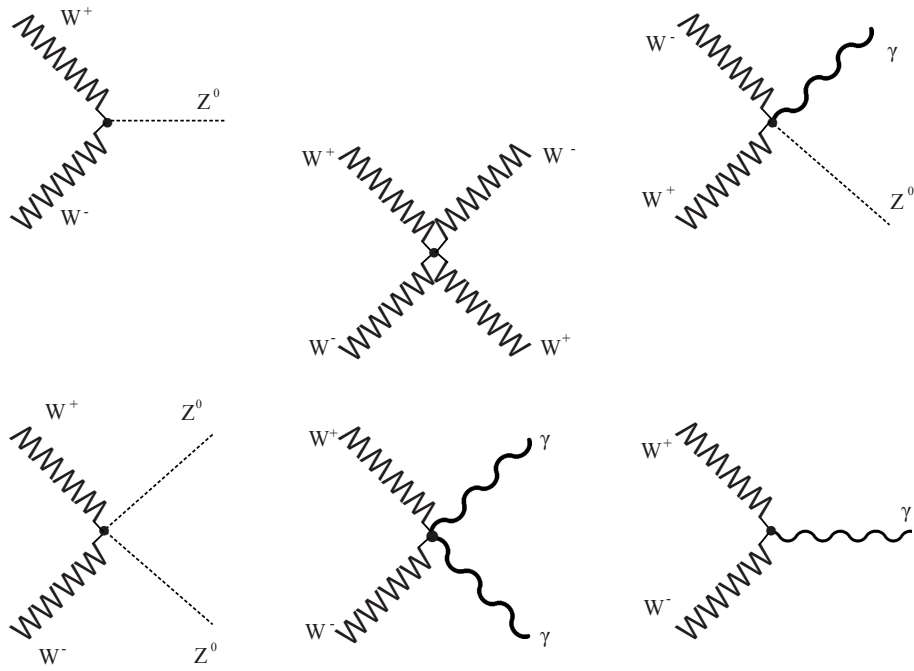


Figure 2.3: Coupling terms between $SU(2)_L \otimes U(1)_Y$ gauge bosons.

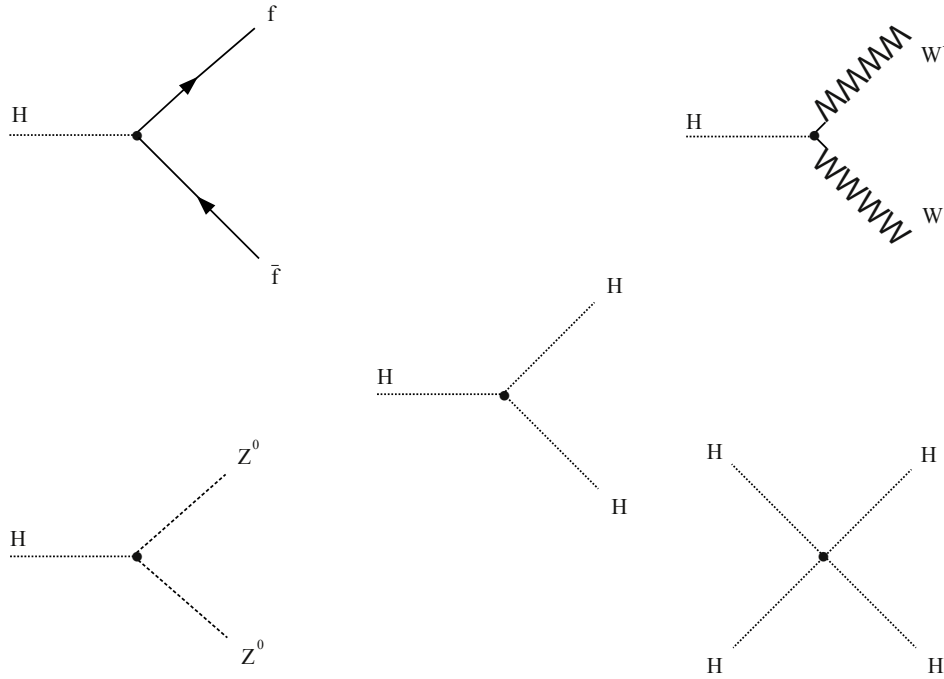


Figure 2.4: Couplings between the scalar field with fermions, $SU(2)_L \otimes U(1)_Y$ massive gauge bosons and selfcouplings.

2.1.2 The $SU(3)_C$ Formalism of Strong Interactions

Quantum Chromodynamics (QCD) is a quantum field theory which describes the interaction between quarks and gluons and together with the electroweak theory, it forms the basis of the Standard Model of elementary particles.

Quarks are fundamental spin-1/2 fermions, whereas gluons are massless spin-1 gauge bosons which mediate the strong interaction.

In a similar fashion to Quantum Electrodynamics (QED) where the generator of the interaction is the electric charge, the charge associated with the strong interaction is called *color* and can take three possible values, symbolically called “red” r , “green” g and “blue” b . Gluons, as the interaction mediators, carry both a unit of color and anticolor and come in eight types.

The color is a hidden parameter, meaning that it can not be directly observed, and, as a consequence, all naturally occurring particles are color singlets. Each quark comes in six flavors, called *up*, *down*, *charm*, *strange*, *top* and *bottom*, and each flavor can be found in one of the three fundamental colors. Anti-quarks carry of course anticolor.

The coupling strength associated with strong interactions - α_s , at small momentum trans-

fer below the QCD scale ($\Lambda_{QCD} \sim 200$ MeV), has relatively high values and a perturbative description cannot be applied. Equivalently the coupling between quarks and gluons (and gluons and gluons) becomes stronger as the separation between them increases, a process known as *confinement*.

Unlike QED where the coupling constant increases with the increase of momentum transfer - q^2 , the fundamental coupling strength (or the color charge) in QCD decreases with increase of q^2 , a fortunate consequence of vacuum polarization, allowing a perturbative description of processes at high momentum transfer. Diagrams with a quark and an antiquark in the loop tend to increase the value of coupling strength where gluon loop diagrams tend to decrease it (see Figure 2.5). The overall result is a decrease of the fundamental coupling with increasing q^2 in accordance with the following relation:

$$\alpha_s(q^2) = \frac{12 \cdot \pi}{(11n - 2f) \cdot \ln(q^2/\Lambda^2)}. \quad (2.25)$$

where n corresponds to the number of colors, f to the number of quark flavors and Λ is a reference energy scale. Consequently at small distances the quarks behave as free particles, an effect known as *asymptotic freedom*.

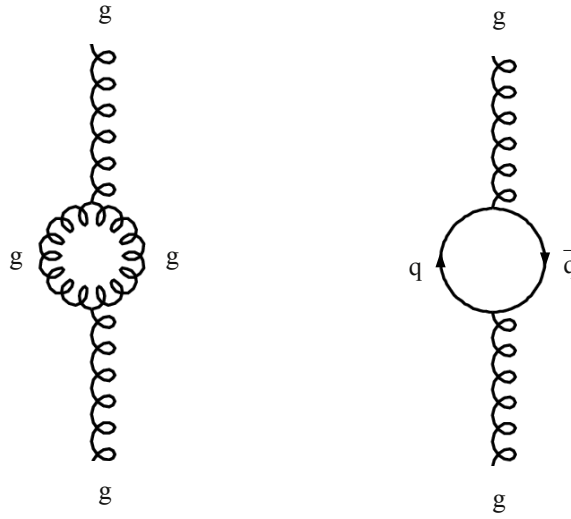


Figure 2.5: Vacuum polarization Feynman diagrams.

Baryons are bound states of three valence quarks where each individual quark carries a distinct unit of color making the combined system color neutral. Mesons are bound states of a colored quark and an anticolored antiquark, the overall color configuration being again color neutral.

Just like electromagnetic and weak interactions, strong processes are described by local gauge invariance and can be described by a symmetry group, namely the $SU(3)_C$ group. By extending the original Yang-Mills model to a Dirac Lagrangian with three equal mass fermions:

$$\mathcal{L} = i\bar{\phi}\gamma^\mu \partial_\mu \phi - m \bar{\phi}\phi, \quad \text{with: } \phi = \begin{pmatrix} \phi_r \\ \phi_g \\ \phi_b \end{pmatrix}. \quad (2.26)$$

and requiring $SU(3)_C$ symmetry for the Lagrangian under a local gauge transformation of the combined wave function:

$$\phi \rightarrow e^{i\theta_k \cdot \lambda_k} \phi, \quad \text{with: } k = 1, \dots, 8. \quad (2.27)$$

the ordinary derivative will have to be replaced by the covariant derivative which will bring in eight massless vector fields A_k^μ , essentially the eight gluons:

$$\mathcal{D}^\mu = \partial^\mu + i\alpha_s \lambda_k \cdot A_k^\mu, \quad (2.28)$$

where the λ_k terms correspond to the eight Gell-Mann matrices. The new local gauge invariant Lagrangian will describe the interaction between a *red*, a *green* and a *blue* quark with eight colored gluons as well as the gluon self interactions:

$$\mathcal{L} = (i\bar{\phi}\gamma^\mu \partial_\mu \phi - m \bar{\phi}\phi) - \frac{1}{16\pi} F_k^{\mu\nu} F_{\mu\nu}^k - j_\mu^k \cdot A_k^\mu. \quad (2.29)$$

where $j_\mu^k = \alpha_s \cdot \bar{\phi}\gamma_\mu \lambda^k \phi$ is interpreted as the quark color current which generates the gluon fields and $F^{\mu\nu}$ is the field strength tensor.

The fundamental QCD interaction terms between quarks and gluons or gluons and gluons are shown in Figure 2.6.

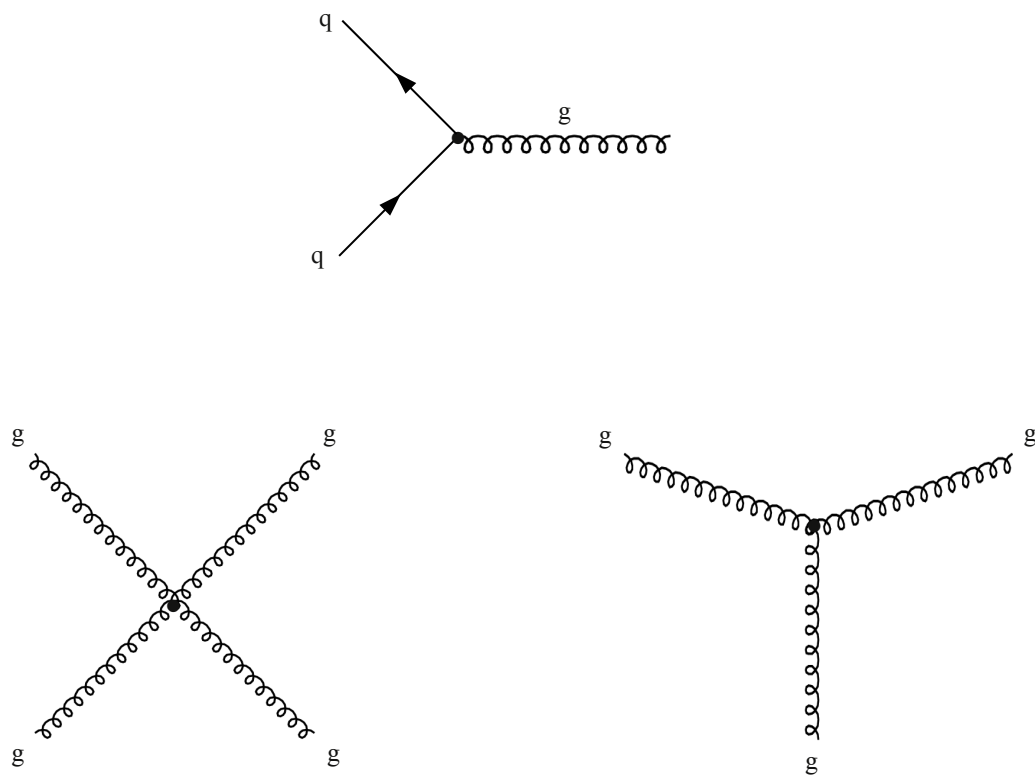


Figure 2.6: The fundamental coupling terms in QCD.

2.1.3 Production of $Z + b\bar{b}$ events

The central subject of this thesis is a measurement of the inclusive cross-section for $Z + b\bar{b}$ production with the ATLAS detector at $\sqrt{s} = 7$ TeV. A precise experimental measurement of the cross-section for such processes would represent a direct test for perturbative QCD predictions (see figure 2.7) as well as an estimate of an important background for the SM scalar boson production in association with a Z boson or the *Minimal Supersymmetric Standard Model* (MSSM) scalar boson production in association with a pair of b -quarks (see figure 2.8).

The Z boson decay channels considered in this analysis are $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$. A theoretical leading order $Z + b\bar{b}$ ² cross-section times branching ratio prediction with the ALPGEN generator³, scaled to next-to-leading order using k-factors⁴ from [19], gives an estimated value of 25.8 pb, resulting in an event yield of approximately 120k for the sum of both leptonic decays of the Z boson, for the $\int \mathcal{L}(t) dt = 4.64 \text{ fb}^{-1}$ integrated luminosity collected during the year 2011.

2.1.4 Production of ZH events

The production and identification of the scalar boson - H in association with a Z^0 (see Figure 2.8) constitutes an important test of the Standard Model electroweak symmetry breaking mechanism. For a low scalar mass up to approximately 130 GeV, the main decay channel is $H \rightarrow b\bar{b}$, resulting thus in a $Z + b\bar{b}$ final state⁵. Direct observation of such events will shed light upon the presence of Yukawa coupling terms, providing thus a mechanism for fermion mass generation within the SM.

A first step in measuring such events begins with a good understanding of the main backgrounds out of which $Z + b\bar{b}$ (see figure 2.7) processes are dominant and irreducible. For a scalar mass of 125.5 GeV, the expected cross-section⁶ at $\sqrt{s} = 7$ TeV is 0.3309 pb [20], resulting in roughly 1500 ZH events produced during the year 2011. At the same mass, the $H \rightarrow b\bar{b}$ branching ratio is approximately 56.9% [20], resulting in a cross-section times branching ratio of 0.168 pb and $\approx 0.8\text{k}$ ZH events in the 2011 data sample.

²Approximately 80% of the total production cross-section at $\sqrt{s} = 7$ TeV for $Zb\bar{b}$ events comes from gluon fusion processes.

³For more details see section 4.2.

⁴The “k” factor is defined as the ratio between the next-to-leading order and leading-order cross-section for a given process.

⁵Though the cross-section for the production of the scalar boson alone, where $H \rightarrow b\bar{b}$, would be significantly higher than the associated production with a gauge boson, the overwhelming background from pure QCD processes such as $g + g \rightarrow b\bar{b}$ would make the measurement highly ineffective.

⁶The cross-section has been computed at next-to-next-to-leading order QCD and next-to-leading order EW.

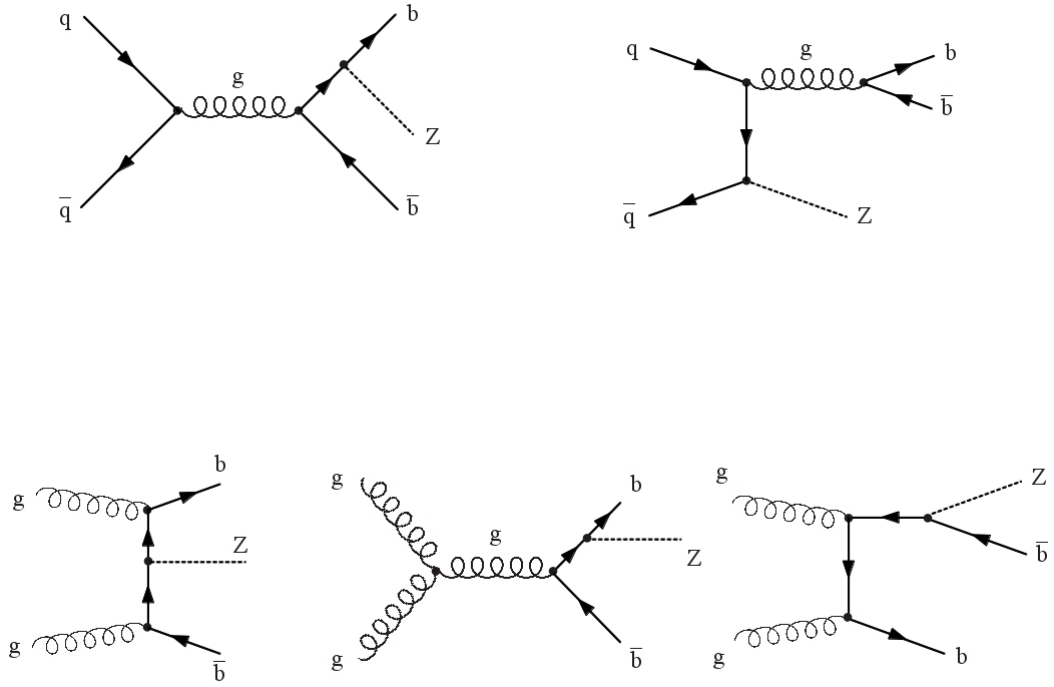


Figure 2.7: Tree level Feynman diagrams for production of $Z + b\bar{b}$ events via gluon splitting (top two diagrams) or gluon fusion (bottom three diagrams) processes. The approximate cross-section contributions for the gluon fusion and the gluon splitting components are 80% and 20% respectively.

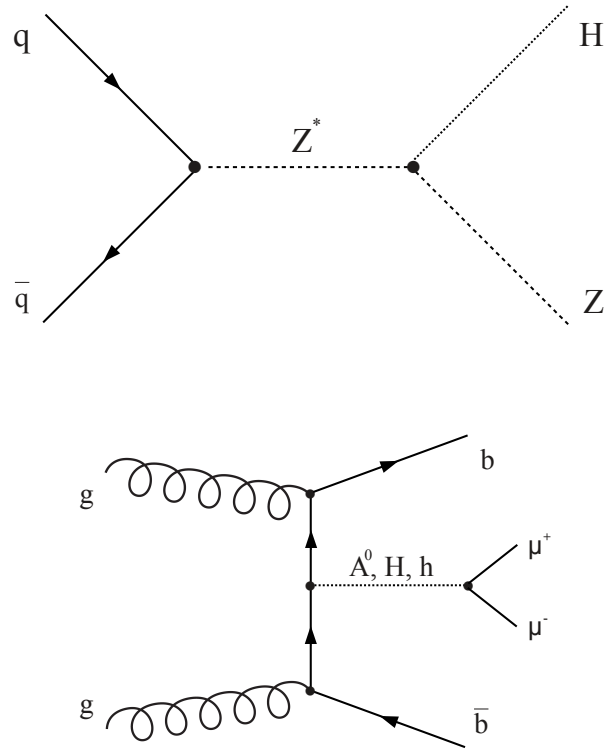


Figure 2.8: Tree level Feynmann diagrams for SM ZH production (top plot) and the MSSM scalar production in association with a pair of $b\bar{b}$ -quarks (bottom plot).

2.2 Experimental Ingredients

Some aspects of the underlying physics occurring in proton-proton collisions at the Large Hadron Collider (LHC) are summarized in this section. The LHC currently holds the man-made world record of center-of-mass collision energy (\sqrt{s}) of 8 TeV. The LHC has been preceded by the successful programme of the Tevatron⁷. The designed maximum center-of-mass energy of 14 TeV is expected to be reached in the year 2015 following the upgrades from 2014. High instantaneous luminosities up to $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, ensured by bunches with lots of particles⁸ above 10^{11} colliding at a frequency of ≈ 11000 Hz per bunch, should create the necessary conditions demanded by many *new physics* scenarios as well as tests for SM processes partially never measured before.

The structure of this section is as follows. Section 2.2.1 will concentrate on the luminosity estimate as a function of beam parameters which is an important input to cross-section measurements.

Protons are composite systems of quarks and gluons, generically called *partons*. An understanding of the interaction probability as function of the momentum fraction of the two colliding partons in events with high- p_T transfer, a process known as *hard scatter*, is thus necessary and will be presented in Section 2.2.2.

Besides the hard scatter, additional low transverse momentum transfer interactions occur between partons inside the same protons or between protons inside the same colliding bunches. This will be covered in Section 2.2.3.

In Section 2.2.4 some models for the hadronization of partons will be described.

2.2.1 Luminosity and Cross-Section Measurements

The expected number of events of a given process depends primarily on two main parameters, the process *cross-section* and the *integrated luminosity* - \mathcal{L} .

The cross-section - σ , is an estimate of the probability for the occurrence of a particular physical process. It is a function of the center-of-mass energy as well as the type of interaction, among other variables. Cross-sections are expressed in units of *barn*⁹.

One of the main parameters at particle colliders is the integrated luminosity, a quantity to which the event yield scales linearly. The *instantaneous luminosity* - $\mathcal{L}(t)$, is defined as

⁷The Tevatron was a $\sqrt{s} = 1.96$ TeV proton-antiproton collider at Fermilab, east side of Batavia, Illinois, USA. One of its greatest achievements was the discovery of the *top*-quark in 1995 by the D0 and CDF experiments.

⁸Proton beams are composed of “clusters” with a large number of protons called *bunches*. The bunch size is confined against the electrostatic repulsion by strong quadrupole magnetic fields at the interaction points.

⁹One barn is equal to 100 fm^2 or 10^{-28} m^2 . More used in particle physics are subunits of the barn such as the *pibarn* 10^{-40} m^2 , the *femtobarn* 10^{-43} m^2 or the *attobarn* 10^{-46} m^2 .

the flux of particles per unit of time and is commonly expressed in units of $\text{cm}^{-2} \text{s}^{-1}$ in accelerator physics. Since the instantaneous luminosity can present large variations over small periods of time and depends quite strongly on the beam parameters, a more useful quantity is the integrated luminosity defined as:

$$\mathcal{L} = \int L(t) dt. \quad (2.30)$$

expressed in units of picobarn^{-1} , femtobarn^{-1} or attobarn^{-1} .

Having an estimate of the integrated luminosity and knowing the cross-section, one can assess the number of expected events in a given process via the simple relation:

$$N^{\text{events}} = \sigma^i \times \mathcal{L}. \quad (2.31)$$

The instantaneous luminosity can be expressed as:

$$L(t) = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y}. \quad (2.32)$$

where f is the collision frequency, n_1 and n_2 stand for the number of particles contained in the two colliding bunches and σ_x , σ_y are the width of the beam profile, assuming gaussian shape, on the x axis - σ_x or on the y axis - σ_y [8].

2.2.2 Parton Distribution Functions

During the collision of two protons at LHC the hard scattering process essentially takes place between two partons which carry only a fraction of the total center-of-mass energy. The probability density of finding the i -th parton inside the proton with the longitudinal momentum fraction - x ¹⁰ at Q^2 scale is called *parton distribution function* - $f_i(x, Q^2)$, usually referred to as PDF. Besides the gluon field and the valence quarks inside the proton, there are also sea-quarks such as u , d , c , s , b (as well as antiquarks) or *gluons* originating in quantum fluctuations inside the proton. At low momentum transfer Q^2 , valence quarks carry most of the momentum fraction of the proton. As the magnitude of the momentum transfer increases, the Bjorken fractions of gluons and sea quarks increase (see Figure 2.9).

A pure theoretical estimation of the PDFs is involving non-perturbative QCD. As a consequence the values of the PDFs are measured at different Q^2 scales and extrapolated to other momentum transfer scales. Once the values of the PDF at the relevant Q^2 scale are

¹⁰The Bjorken momentum fraction x can be interpreted as the ratio between the parton's longitudinal momentum p_i and the longitudinal momentum of the proton p : $x_i = p_i/p$.

known the cross-section for a given process (e.g. production of particles a and b) at LHC can be computed:

$$\sigma = \sum_{i,j} \int \int f_i(x_i, Q^2) dx_i \cdot f_j(x_j, Q^2) dx_j \cdot \sigma_{ij \rightarrow ab} \quad (2.33)$$

where i and j symbolize the colliding partons inside the incoming protons.

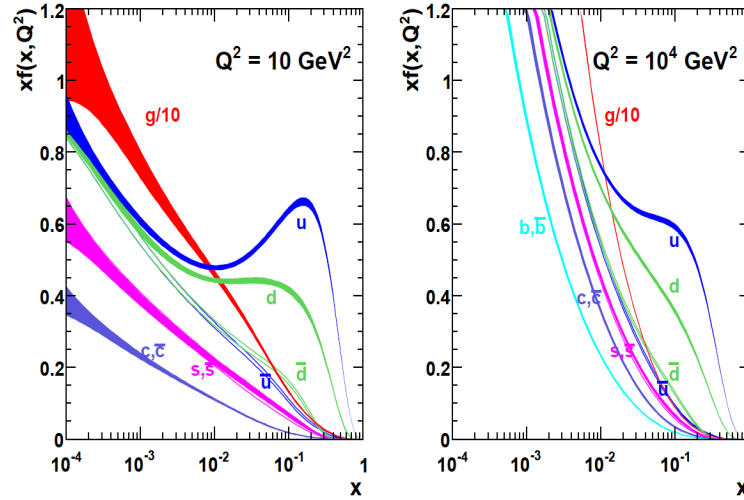


Figure 2.9: The MSTW2008 PDF parametrization at NLO at $Q^2 = 10 \text{ GeV}^2$ - left and $Q^2 = 10^4 \text{ GeV}^2$ - right [11].

2.2.3 Pileup, Multiparton Interactions and the Underlying Event

During the bunch crossings inside the collider, besides the high- p_T hard scatter process, there is an additional low transverse momentum component originating in multiple proton-proton inelastic scattering events, taking place at much higher rates than the hard core component. Though the transverse momentum transfer in such collisions is relatively small, they will produce additional low- p_T hadrons which can deposit energy inside calorimeter and the tracker systems. One needs to consider such events since they are characterized by relatively high cross-sections and can alter the energy scale and resolution for physically interesting objects such as high- p_T electrons, muons or jets. Such soft inelastic proton-proton collisions are called *pileup*¹¹. The average number of pileup

¹¹There are actually two different types of pileup. The proton-proton collisions from the same bunch crossing are called *in-time pileup*. In events with low $\langle \mu \rangle$ and relatively long bunch crossing separation, it

interactions per bunch crossing, $\langle\mu\rangle$, during the 2011 data taking was measured around 9-10 as shown in figure 2.10.

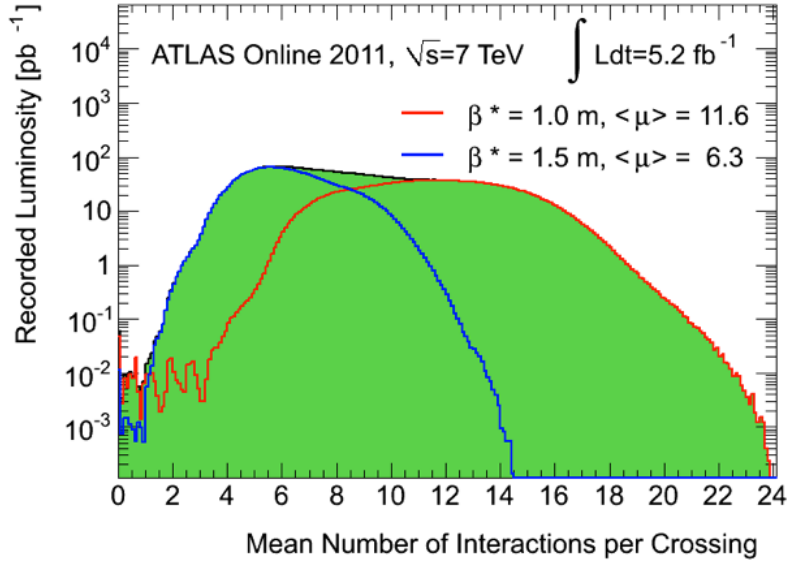


Figure 2.10: The distribution of the average number of interactions per bunch crossing weighted by integrated luminosity specific to 2011 data. The curves correspond to data collected prior (blue) and after (red) the 2011 September Technical Stop. The β^* parameter corresponds to the distance from the nominal interaction point to the point where the beam thickness is twice as large as at the nominal interaction point [12].

Multiple parton interactions (MPI) describe additional parton-parton scatterings besides the main hard interaction from the same proton-proton collision. In a way, the process resembles pileup, only in this case the scatter takes place between different partons from the same colliding protons whereas for pileup the scatter occurs between different protons from the same incoming bunches.

By taking out the hard component from the inclusive scattering process the remainder constitutes the *underlying event* component which sums up pileup, multiple parton interactions, elastic scatter, as well as single and double diffractive interactions.

can be measured by the number of reconstructed vertices. This was the case for data taken during 2010. As the bunch time separation decreases, soft pileup remnants from consecutive bunches start to overlap. This happens because the detector response is slower than the time between two successive bunch crossings. Such effects, called *out-of-time pileup* were common in most of 2011 Data.

2.2.4 Parton Shower and Hadronization Models

The *parton shower* is a phenomenological model of the QCD radiation which begins virtually immediately after the hard interaction took place and consists, in a first stage, in the perturbative emission of gluons from the partons resulted in the hard scatter which subsequently split into quark-antiquark pairs. The resulting $q\bar{q}$ pairs will radiate further gluons and the process continues as long as a perturbative QCD description can be applied, usually chosen up to a phenomenological infrared cutoff scale Q_{had} , i.e. the *hadronization scale*.

Hadronization models describe the non-perturbative QCD process, subsequent to the parton shower, corresponding to the conglomeration of colored quarks into color singlet bound states such as mesons and baryons. The scale at which the parton shower process ends and the hadronization begins, Q_{had} , is a model dependent parameter usually chosen near the Λ_{QCD} scale.

A pure and complete theoretical description of hadronization involves a non-perturbative treatment of systems with usually high parton multiplicities, raising, as a result, numerous practical and theoretical difficulties. A more successful approach is realized in phenomenological models such as the *Cluster Model* or the *Lund String Fragmentation Model*, briefly discussed in the following.

The central idea of the Lund String Fragmentation Model [8] assumes the presence of a strongly interconnecting color-dipole tube-shaped field (a string) between quark-antiquark pairs generated in the hard scattering. While the distance between the partons grows, the strong interaction potential increases linearly¹², an effect known as “*linear confinement*”. When the potential energy accumulated in the string exceeds a critical value, it breaks down, forming bound hadron states such as mesons or baryons¹³. The process continues until all the available potential energy is converted into hadrons. The string hadronization model is used by MonteCarlo generators such as PYTHIA [14].

A rather different hadronization approach is addressed by the Cluster Model [8] which conjectures the formation of local color singlet clusters composed of $q\bar{q}$ pairs at the end of the parton shower (“*preconfinement*”). Each cluster, produced via non-perturbative $g \rightarrow q\bar{q}$ splitting of all gluons produced in the parton shower, will decay either into a single hadron¹⁴ - if its invariant mass is below some phenomenological cutoff¹⁵, or, it can decay into two sub-clusters and the process continues until all sub-clusters have decayed into hadrons. The Cluster Hadronization Model is used by the HERWIG [15] or SHERPA [18] MonteCarlo generators.

¹²The strong potential between $q\bar{q}$ pair can roughly be approximated as: $V(r) \approx k \cdot r$ with $k \approx 1 \text{ GeV/fm}$.

¹³Mesonic systems form when the string splits into quark-antiquark pairs. On the other hand the rupture of the string into diquark-diantiquark pairs will lead to baryonic states.

¹⁴The eventual energy excess being transferred to the neighboring clusters.

¹⁵Usually taken around 4 GeV

3 Experimental Setup

This chapter is dedicated to some of the technical details of the complex data-taking machinery of the LHC accelerator and the ATLAS detector.

First, a description of the LHC proton synchrotron will be given in Section 3.1 followed in Section 3.2 by a general overview of the ATLAS detector.

Subsystems of the ATLAS detector such as the Inner Detector, the Electromagnetic and Hadronic Calorimeters, the Muon Spectrometer as well as the Magnet and Trigger systems and the forward detectors are discussed in Sections 3.3 - 3.7. A short description of the Worldwide LHC Computing Grid and MonteCarlo generators will end the chapter.

3.1 The Large Hadron Collider

On September 10th, 2008, for the first time, proton beams circled successfully the LHC ring. Nine days later a regretful incident ¹, postponed the long awaited start of data taking by one more year.

The first collisions after the 2008 incident, occurred on November 20th, 2009. Ten days later the world record of $\sqrt{s} = 2.36$ TeV center-of-mass collider energy is reached, opening the gate to a higher energy frontier and possibly to a new physics chapter.

The Large Hadron Collider is located about 100 m underground CERN's ² site on the swiss-french border, north-west of Geneva, Switzerland. Its tunnel, 26.7 km in circumference, hosted the LEP accelerator ³ between 1989 and 2000. The LHC has been designed as a high-energy investigation tool of the Standard Model, signs of physics beyond the

¹An electrical connection failure caused a quench in the magnet dipole system in sectors 3 and 4. As a result approximately 6 tons of liquid helium were poured into the LHC tunnel. Harsher precaution measures had to be taken which redefined the initial LHC program.

²*Conseil Européenne pour la Recherche Nucléaire* initially. Later the name was changed to *European Organization for Nuclear Research*.

³The Large Electron Positron Collider (LEP) collided electrons with positrons up to $\sqrt{s} = 209$ GeV and was designed for precise measurements of electroweak parameters. LEP remains until today the most powerful lepton accelerator ever constructed.

Standard Model as well as the discovery (or exclusion) of the long sought electroweak symmetry breaking scalar particle. Its main detectors, ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), LHCb (Large Hadron Collider beauty) and ALICE (A Large Ion Collider Experiment) ⁴ are placed at the four interaction points.

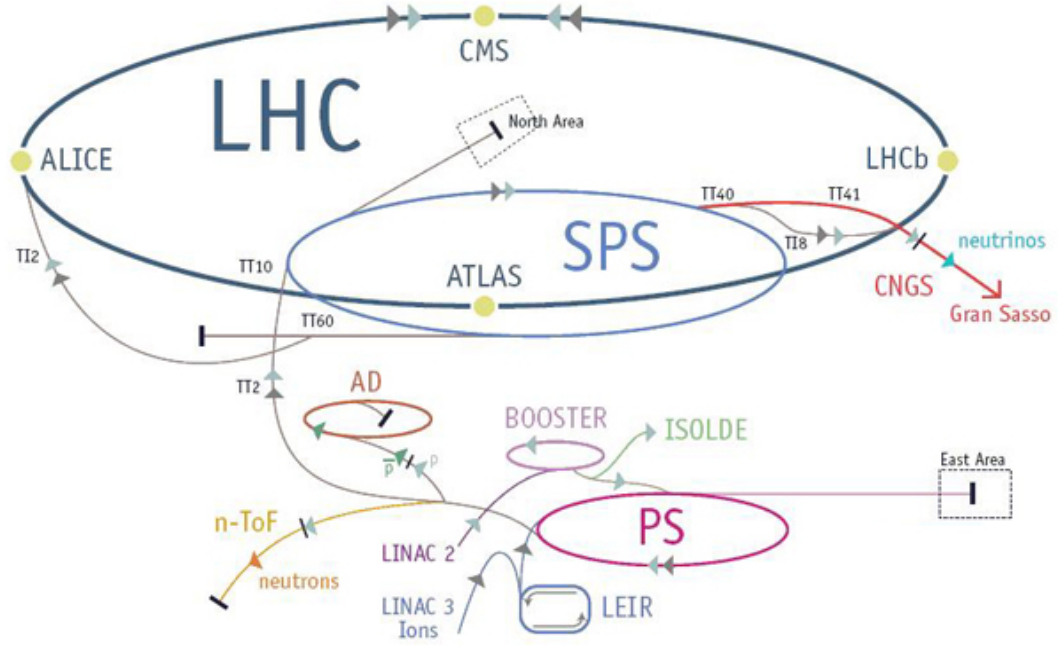


Figure 3.1: The path followed by particles prior to collisions. For experiments such as ATLAS, CMS or LHCb the proton's path is: LINAC 2 → PSB (BOOSTER) → PS → SPS → LHC. For ALICE the ion lead's route is: LINAC 3 → LEIR → PS → SPS → LHC [21].

ATLAS and CMS are general-purpose experiments, their research spectrum including searches for supersymmetry, large extra dimensions, exotic particles as well as new sources of CP-violation, the discovery or exclusion of the scalar particle, top physics, QCD as well as many other SM processes.

The LHCb experiment is measuring CP-violation within the SM as well as searching for signs of additional CP-violation beyond the SM, the electroweak symmetry breaking scalar boson and particularly hints of new physics occurring in bottom-hadron decays.

Unlike the other three main experiments at the LHC, ALICE is a heavy ion experiment measuring parameters of the quark-gluon plasma resulted in $Pb-Pb$ collisions up

⁴The LHC also hosts three much smaller sized experiments - TOTEM (TOTal Elastic and diffractive cross-section Measurement), LHCf (Large Hadron Collider forward) and MoEDAL (Monopole and Exotics Detector at the LHC) having a narrower research area.

to $\sqrt{s} = 5.52$ TeV per nucleon.

Before entering the LHC ring, the protons' energy is increased in a number of successive steps (see Figure 3.1). The route starts with LINAC 2⁵ which accelerates protons to 50 MeV before injecting them into the Proton Synchrotron Booster (PBS) where their energy is further increased to 1.6 GeV. After the PSB the protons enter the Proton Synchrotron (PS) where they are accelerated to 26 GeV and from there they are dumped into Super Proton Synchrotron (SPS)⁶. At the end of the SPS cycle the protons energy is 450 GeV. Finally the SPS injects protons into the LHC ring where the final step of acceleration takes place. In 2011 the maximum center-of-mass proton-proton collision energy was 7 TeV and in 2012 the energy was increased to 8 TeV.

3.2 The ATLAS Detector

ATLAS [22] is one of the two multi-purpose detectors at the LHC. It presents a forward-backward geometric symmetry having its components arranged cylindrically around the beam-pipe. The center of the detector is placed at the nominal interaction point, the region where the two incoming beams intersect. ATLAS measures approximately 25 m in diameter, 45 m in length and weighs more than 7000 tonnes, making it one of the largest particle detectors ever constructed. The ATLAS detector uses a right handed coordinate system (\vec{r}, θ, ϕ) having its origin at the nominal interaction point. The x and y axis point to the center of the LHC and upwards respectively. The beam direction at the collision point gives the z axis direction. Angular coordinates are measured in terms of the azimuthal angle - ϕ ⁷ and the pseudorapidity - η ⁸. Distances, measured in the η - ϕ , are computed as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The success of the ATLAS program depends directly on the efficiency of ATLAS detector which itself relies on the performance of each of its components. Following the route of particles produced at the collision point, the first subsystem encountered is the Inner Detector surrounded by a superconducting solenoid magnetic system. Next come the Electromagnetic and Hadronic Calorimeters and finally the last and largest sub-system is Muon Spectrometer (see figure 3.2).

⁵LINAC 2 is a linear accelerator constructed between 1973 and 1978 in order to meet the technical requirements of the "new" Proton Synchrotron Booster (finalized 1972) uncovered by its predecessor - LINAC 1.

⁶The SPS has provided the first direct signals of the existence of the Z^0 and W^\pm vector bosons in the 1984 by colliding protons with antiprotons.

⁷The azimuth is defined as the angle made by the projection of the \vec{r} vector in the x - y plane and the x axis.

⁸The pseudorapidity is defined as: $\eta = -\log(\tan\theta/2)$, where θ is the angle between \vec{r} and the z axis.

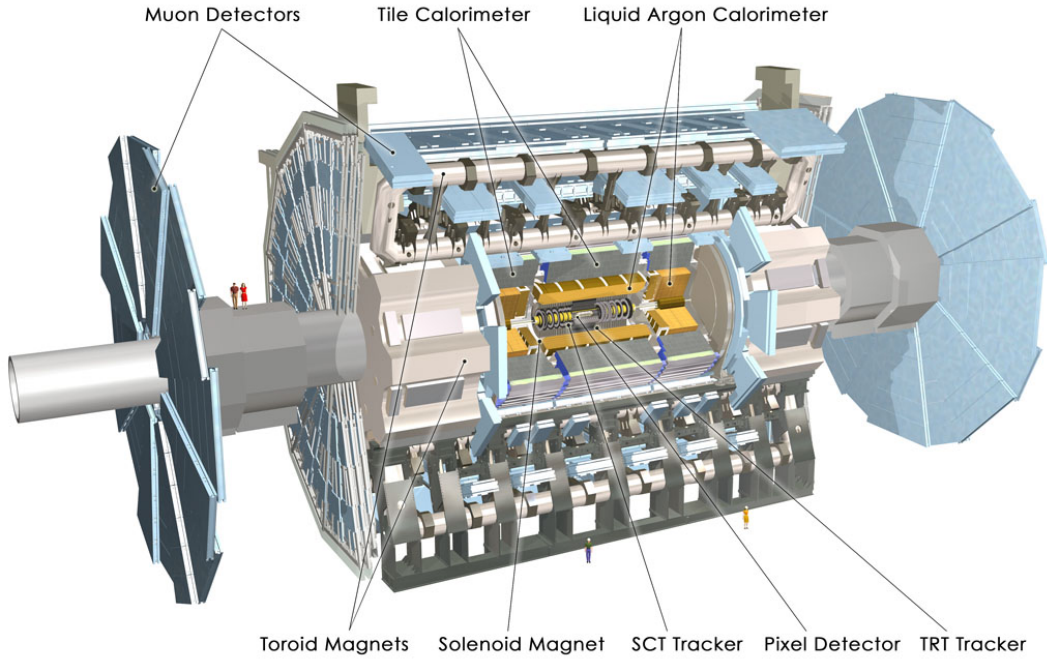


Figure 3.2: Overview of the components of the ATLAS detector [22].

3.3 The Inner Detector

Precise tracking, charge identification, primary and secondary vertex reconstruction as well as high momentum resolution⁹ are just a few mandatory requirements in the high particle multiplicity environment present in the vicinity of the nominal interaction point. Such prerequisites are successfully met by the ATLAS Inner Detector (ID). The ID measures 2.1 meters in diameter and extends approximately 6.2 m in length, covering a pseudorapidity range from -2.5 to +2.5. In order to minimise the energy loss during the passage of charged particles, the amount of material crossed varies between 0.4 and maximum 2 radiation lengths. The 2 T magnetic field generated by a surrounding solenoid allows charge sign and momentum measurements by magnetic deflection of charged tracks. The main components of the ID are the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT), as shown in figure 3.3.

⁹The Inner Detector yields an average transverse momentum resolution of $\Delta p_T/p_T = 0.05\% \times p_T(\text{GeV}) \oplus 1\%$.

The **Pixel Detector** is the sub-detector system closest to the beam pipe, being designed, together with the SCT, for accurate 3-dimensional track and vertex reconstruction. Consisting of three cylindrical layers with a decreasing granularity around the beam line in the barrel region and three disk layers in each of the end-caps, it covers regions up to $|\eta| = 2.5$. The high granularity of the Pixel detector is assured by 1744 silicon-pixel sensors, 90% of them measuring $50\ \mu\text{m}$ in $r\text{-}\phi$ direction and $400\ \mu\text{m}$ in z ($50\ \mu\text{m}$ in $r\text{-}\phi$ and $600\ \mu\text{m}$ in z for the remaining 10 %), reaching a typical sensor resolution of about $10\ \mu\text{m}$ in $r\text{-}\phi$ and $120\ \mu\text{m}$ in z . In total there are more than 80 million readout channels for the entire Pixel system. Precise reconstruction of secondary vertices originating in decays of b -hadrons is an important input to b -tagging algorithms which is provided by the Pixel together with the Semiconductor Tracker. Strong rejection of electrons originating from photon conversions is provided by the innermost Pixel layer (B-layer), by requiring track hits in this region [22].

Following the Pixel Detector, the **Semiconductor Tracker** consists of four cylindrically positioned double-layered stereo micro-strip sensors in the barrel region and nine disks of micro-strips in the two end-cap wheels, counting together 4088 modules and a total of 6.3 million readout channels. The mean strip-pitch is $80\ \mu\text{m}$ and the relative inclination between the two sets of micro-strips in the barrel is approximately 2.3 angular degrees. Track coordinates in $r\text{-}\phi$ are measured by 6.4 cm long rectangular silicon stereo micro-strips where one set of strips has been arranged parallel to the beam line. The resolution per module is approximately $17\ \mu\text{m}$ in $r\text{-}\phi$ and $580\ \mu\text{m}$ in z . The trapezoidal end-cap sensors yield a similar accuracy per module to those from the barrel and have been arranged in two sets ¹⁰ [22].

The last and most massive component of the Inner Detector is the **Transition Radiation Tracker** system, covering a radial region from 55 cm to 108 cm from the beam pipe. Approximately 350000 readout channels from more than 175000 drift tubes are providing additional $r\text{-}\phi$ track information ¹¹ and cover a $|\eta| < 2$ region. The active area of TRT tubes is provided by a xenon-based gas mixture (70%Xe : 27%CO₂ : 3%O₂) producing charged clusters in regions where the charged tracks have interacted with the gas molecules. The positive charged pulse is picked up by a gold-plated central wolfram wire while the negative charge is collected by the tube's walls. The average precision per straw is $\approx 130\ \mu\text{m}$. Filling the space between the tubes, a system of layers with varying dielectric constant stimulates the production of transition radiation during the highly relativistic charged particles passing. In the barrel region the straws are placed parallel to the beam line and measure around 1.4 m in length while in the end-cap wheels they are set radially and have an average length of 37 cm.

¹⁰A first set of strips is placed radially on the wheels while a second set, composed of stereo micro-strips has been inclined with respect to the first by an angle of 2.3 degrees.

¹¹Unlike the Pixel or the SCT, the TRT only measures $r\text{-}\phi$ track coordinates.

The mean number of hits per track with $0.5 \text{ GeV} < p_T < 100 \text{ GeV}$ ranges from 22 in the barrel-end-cap transition region to 35-36 outside this region. Transition radiation induced inside the TRT medium provides a powerful tool to disentangle electrons and charged mesons [22].

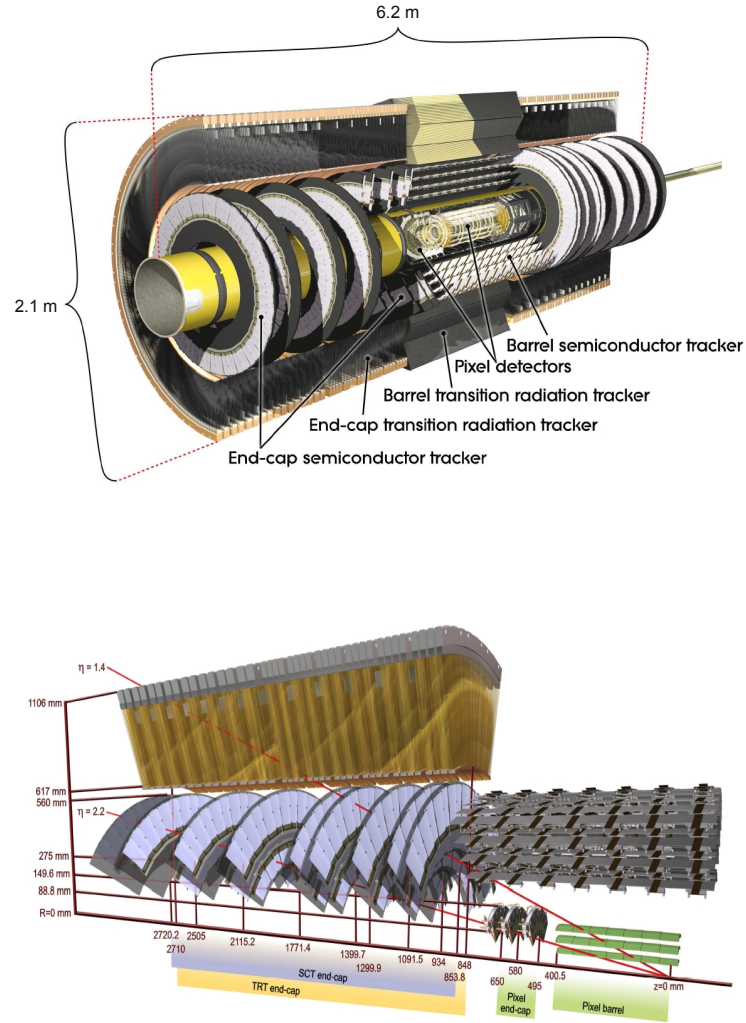


Figure 3.3: A detailed overview of the internal structure of the Inner Detector is shown in the top figure. A cross-section sketch of the ID barrel is shown in the bottom figure [22].

3.4 The Solenoid Magnet

A superconducting 2 T **Central Solenoid Magnet** (see figure 3.4) is located between the Inner Detector and the Electromagnetic Calorimeter. It has the role of bending the trajectory of charged particles and provide the possibility of measuring transverse momentum components. With a diameter of 2.494 m and a length of 5.282 m, the central solenoid uses a NbTi superconductor alloy cooled at 4.2 K. Weighing 5.7 tons, the material of the superconducting magnet has been minimized in order to reduce interactions with particles emerging from the interaction point [27, 22].

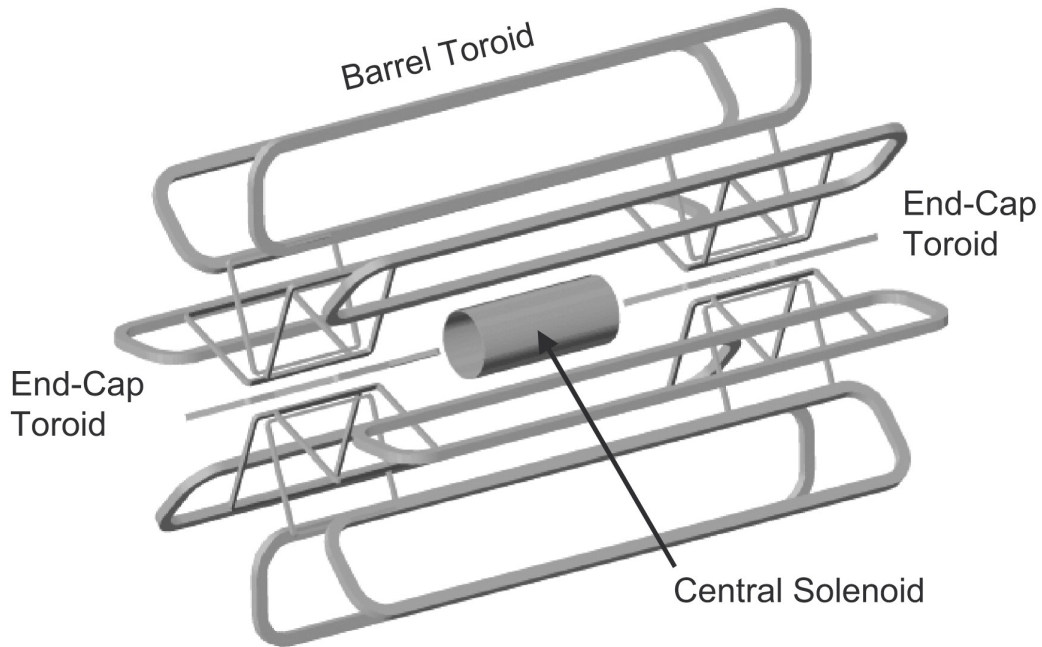


Figure 3.4: ATLAS Magnetic System [27].

3.5 The Calorimeter System

Precise energy and direction measurements of physical objects such as e^\pm , γ , jets and Missing Transverse Energy (MET)¹² as well as triggering¹³ are covered over a wide pseudorapidity range by the Calorimeter System, composed of the Electromagnetic and Hadronic Calorimeters (see Figure 3.5).

The Electromagnetic Calorimeter

Electromagnetic showers produced by the absorption of electrons, positrons and gamma rays¹⁴ are accurately measured via the full and symmetric ϕ coverage of the sampling Electromagnetic Calorimeter (EC). High reconstruction, identification and good energy resolution due to the fine granularity structure of the EM provide an efficient reconstruction of prompt electrons such as those from Z boson decays with relatively good energy resolution: $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\text{GeV}}} \oplus 0.7\%$ [22].

The EC is structured in a central barrel module ($|\eta| < 1.475$) and two double end-cap wheels ($1.375 < |\eta| < 3.2$). The active medium is liquid argon kept at a constant temperature by a cryostat system while the absorber role is played by accordeon-shaped lead plates. A similar accordeon geometry is also used by the kapton electrode system. Three layers with varying granularity in the barrel region give additional particle direction estimates independent of the Inner Detector (first two layers) as well as hadronic-like objects rejection (third layer). The first barrel layer, being closest to the beam line has the finest granularity - $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$ in regions of small pseudorapidity ($|\eta| < 1.4$) and yields high discriminant power between γ or e^\pm and charged mesons. The second layer, with a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, is the region where most of the energy is deposited. Finally the third layer, with the coarsest granularity $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ gives an η coverage up to 1.35. The transition region between the barrel and the end-caps $1.37 < \eta < 1.52$ has been shown to yield poorer reconstruction and identification efficiencies due to the larger amount of detector material placed in front of the first active calorimeter layer. With a thickness of approximately 25 radiation lengths, the EC guarantees a full absorption of electrons and gamma particles [22].

¹²Missing Transverse Energy - E_T^{miss} is defined as the negative algebraic sum of all calorimeter cell-energy deposits plus the energy of all muon tracks measured in the Muon Spectrometer.

¹³The Calorimeter system is seeding the Level-1 trigger via regions of interest.

¹⁴Jets as well as taus or muons can lose a fraction of their energy in the EM system.

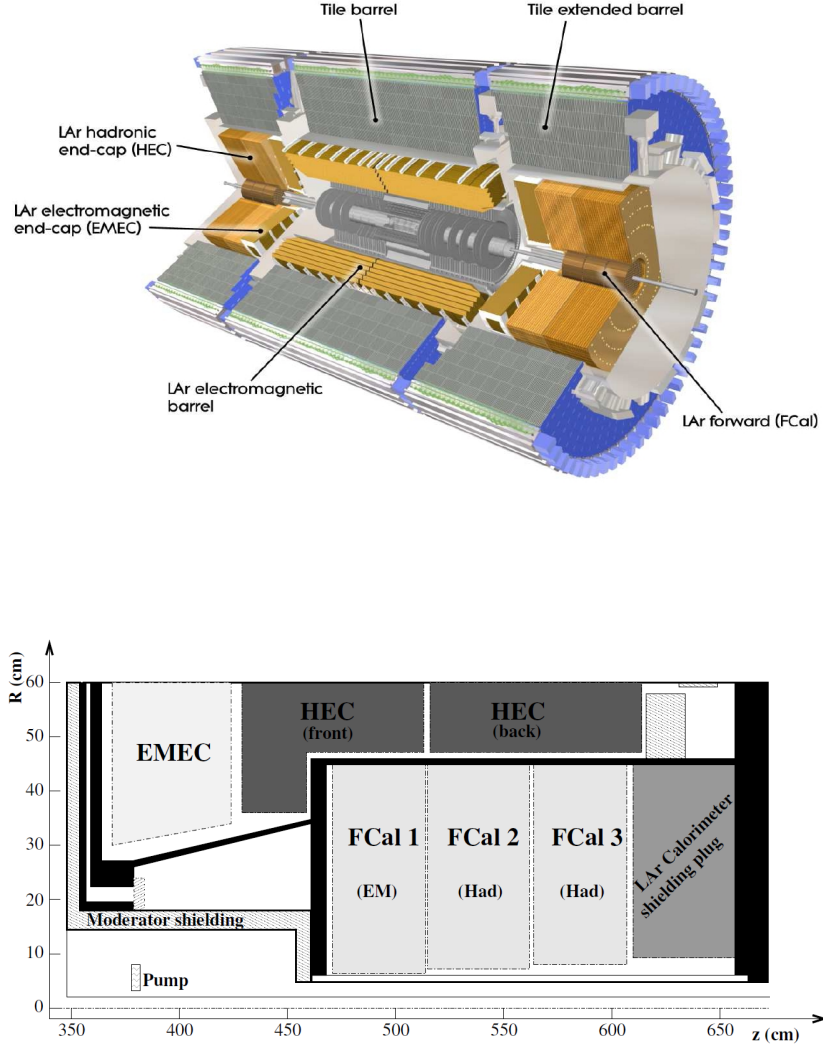


Figure 3.5: Top figure shows a general view of the Calorimeter System of the ATLAS detector [23]. The end-cap and forward calorimeter structures can be seen in the bottom figure [22].

The Hadronic Calorimeter

Designed to measure hadronic-jet induced shower parameters, the Hadronic Calorimeter gives an average energy resolution of $\frac{\sigma_E}{E} = \frac{56\%}{\sqrt{E/\text{GeV}}} \oplus 5.5\%$ [22].

A pseudorapidity region up to $|\eta| < 1.0$ is covered by the central barrel **Tile Calorimeter** while two auxiliary bodies extend from $0.8 < |\eta| < 1.7$. Hadronic showers are produced inside the active iron and scintillating tiles while absorption is ensured by steel based absorbers. The Tile calorimeter extends from 2.28 m up to 4.25 m from the beam pipe¹⁵ and presents a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ granularity. Two sets of double-layered wheels placed at each side of the end-caps provide an $1.5 < |\eta| < 3.2$ coverage and make the body of the **Hadronic End-cap Calorimeter** (HEC) system. The active material is liquid argon while the absorption is provided by copper-based materials [22].

LAr Forward Calorimeter

The LAr Forward Calorimeter (FCal) ensures η coverage from 3.1 to 4.9 in the forward regions. Each forward side consists of three modules (see bottom figure from 3.5) measuring together approximately 10 interaction lengths in depth. The active material for all three modules is liquid argon based. Designed for measurements of electromagnetic-interacting objects, the first module presents copper-based absorbers while the second and third are wolfram-based and measure cluster parameters from hadronic processes [22].

3.6 The Toroid Magnet

The **Toroid Magnetic System** consists of a central barrel and two end-cap bodies, each with an eight-coil structure as shown in figure 3.4. The toroid provides an essential input to the Muon Spectrometer by bending muon tracks in the η direction. It weights about 830 tonnes, measures approximately 25 m in length and 10 m in diameter and yields an average 0.5 T magnetic field. In order to reduce muon scattering, the superconducting toroid's core is air-based [22].

¹⁵Yielding an average of 10 nuclear interaction lengths for relativistic hadrons at $\eta = 0$.

3.7 The Muon Spectrometer

Constructed for triggering and to give independent precise muon track measurements, the large Muon Spectrometer is the outermost detector layer of ATLAS. High momentum resolution¹⁶ and track reconstruction are provided via Monitored Drift Tubes and Cathode Strip Chambers over a wide pseudorapidity range while an independent trigger system is provided by Resistive Plate Chambers and Thin Gap Chambers technologies (see figure 3.6).

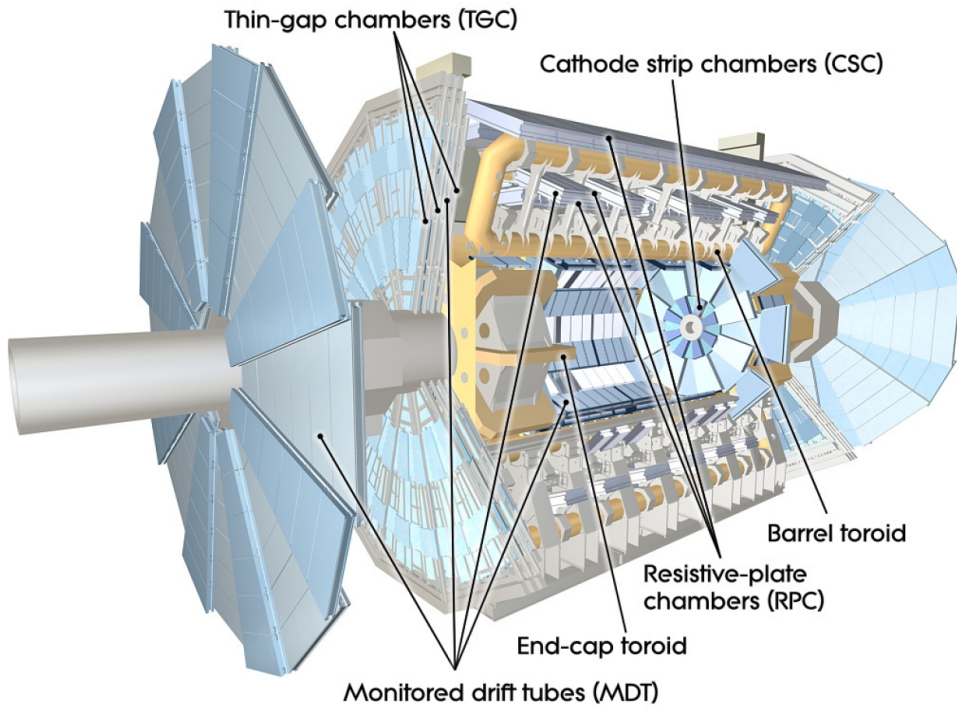


Figure 3.6: Overview of the Muon Spectrometer and its four subsystems [26].

The Muon Spectrometer Trigger System

Three concentric cylindrical layers of **Resistive Plate Chambers** (RPC) with a full ϕ coverage provide the trigger system of the Muon Spectrometer in the barrel region. The RPC system stretches up to $|\eta| < 1.05$. The inner two layers are designed for low p_T muon triggering with three thresholds between 6 GeV and 9 GeV while the outermost layer in conjunction with the inner two, rises the threshold up to 35 GeV in three additional steps (see figure 3.7). Approximately 544 chambers with 360000 readout channels give

¹⁶The muon resolution ranges from 1-2 % for $p_T \sim 300$ GeV up to 10 % for $p_T \sim 1$ TeV [22].

sufficient spatial (10 mm) and time resolution (1.5 ps) for triggering. Each RPC chamber is composed of two identical units and each unit has two gas containers with a thickness of 2 mm (gap) which provide the active medium. The gas is a mixture of $C_2H_2F_4$ - 94.7 %, $IsoC_4H_{10}$ - 5 % and SF_6 - 0.3 % and is closed by two high-voltage bakelite resistive plates. During the charged tracks crossing, the gas produces an avalanche signal which is picked up via two independent layers of strips with an η - ϕ arrangement, placed on opposite sides of the plates. Due to its η - ϕ strip geometry, RPC chambers provide additional second coordinate (ϕ) track information[22, 24, 25].

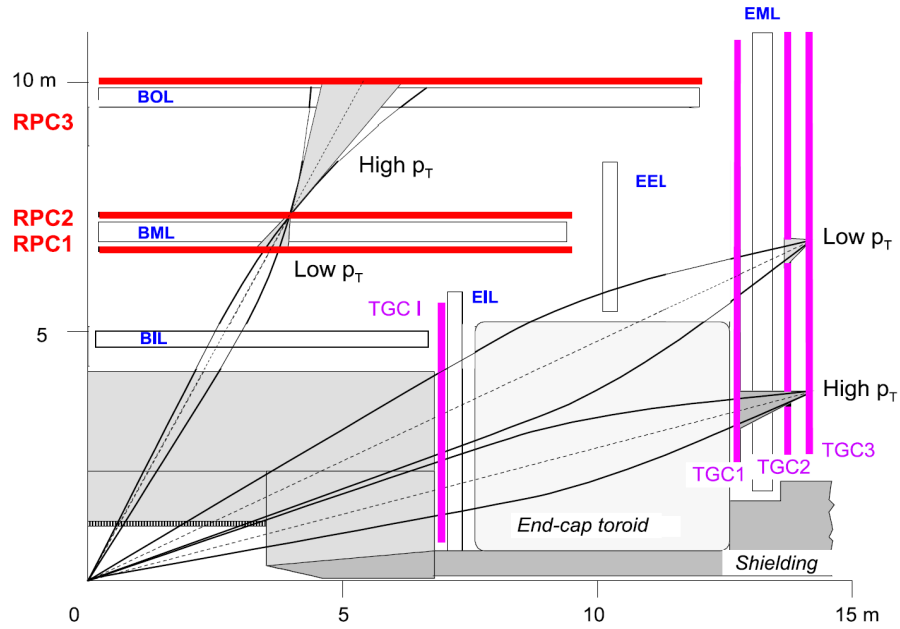


Figure 3.7: View of the barrel and end-cap Trigger System [22].

For the high rate environment present in the end-cap regions, a finer granularity technology has been chosen for the trigger system, namely **Thin Gap Chambers** (TGC). There are four wheels of TGC chambers on each end-cap side as shown in 3.7, TGC 1 having the largest η coverage (up to 2.7). A track deviation from a straight line caused by a toroid's magnetic field is measured by the difference between the reconstructed segment inside different chambers (TGC2&3 or TGC1&2&3) and the track segment between the reference TGC3 layer ("pivot layer") and the interaction point. Since the TGC 2 and TGC 3 wheels are closest to each other, they will trigger on low p_T tracks while TGC1, TGC2 and TGC3 together are used in high p_T triggering. The trigger end-cap system is limited

to $|\eta| < 2.4$, this being the TGC2 and TGC3 wheels coverage ¹⁷.

The TGCs are multi-wire proportional chambers with a combination of CO₂ (55 %) + n – C₅H₁₂ (45 %) as active material. The electric signal is collected by η wires and ϕ pick-up strips. Similar to RPCs, TGCs also provide second coordinate information [22, 24, 25].

The Muon Spectrometer Precision System

The **Monitored Drift Tubes** system (MDT) has the largest spatial extension being ordered in three concentric layers around the beam pipe in the barrel ($|\eta| < 1.3$) and three end-cap wheels ($1.3 < |\eta| < 2.4$). With more than 1088 chambers and 339000 readout channels, MDTs were built to yield excellent tracking, independent of the Inner Detector. A typical MDT chamber has a multi-layer inner structure with each layer containing up to four rows of cathode tubes. With a thickness of $\approx 400 \mu\text{m}$, the walls of a MDT tube are made of Al and function as cathode. A central wire with a diameter of $50 \mu\text{m}$, is used as the anode. The operating gas is a combination of 93% Ar and 7% CO₂ kept at a constant 300000 Pa pressure. Charged clusters created by the muon's interactions with the gas, will drift towards the anode and cathode, being accelerated by a 3080 V potential, resulting in an enhancement of the initial signal via charge-induced avalanches. With a maximum drift time of 700 ns and $80 \mu\text{m}$ resolution, an average tube yields a gas gain of about 20k. The typical chamber resolution is approximately $35 \mu\text{m}$. MDTs provide only precision tracking in the bending (η) direction.

The alignment of MDT chambers is provided by an optical sensor system with a precisions a few mm [22, 24, 25].

Being closest to the beam pipe, the **Cathode Strip Chambers** system (CSC) covers a region from $2.0 < |\eta| < 2.7$ and provides precision tracking via two forward wheels mounted on each side of the barrel. Similarly to TGCs, CSC chambers have a multi-wire proportional internal structure. The anode wires are distributed radially (perpendicular to the beam line) while cathode strips are mounted in the $\eta - \phi$ plane such that muon tracks are reconstructed by the interpolation of the induced charge in the grid strip layers. The active medium is a combination of 80 % Ar and 20 % CO₂ with a gas gap of 5 mm and a gain of 60k. The 32 chambers (16 in each wheel) provide a total of 31000 readout channels. In contrast to MDTs, the CSC chambers measure track coordinates in both η and ϕ [22, 24, 25].

¹⁷TGC I has an $1.05 < |\eta| < 1.92$ extension. TGC1 extends from 1.05 to 2.7 in η while TGC2 and TGC3 cover a region between $1.05 < |\eta| < 2.4$.

3.8 Trigger and Data Acquisition

The ATLAS Trigger system has been designed to deal with the high event rate ($\approx 4 \times 10^7$ Hz) and select interesting physics events in real time. Decisions taken in parallel processing via three successive trigger levels - **LVL1**, **L2** and **Event Filter**, reduce the event rate down to a few hundred Hertz, at which point the data storage process begins. The average event size is approximately 1.5 MB, resulting in a storage speed of roughly ≈ 600 MB/s (see Figure 3.8).

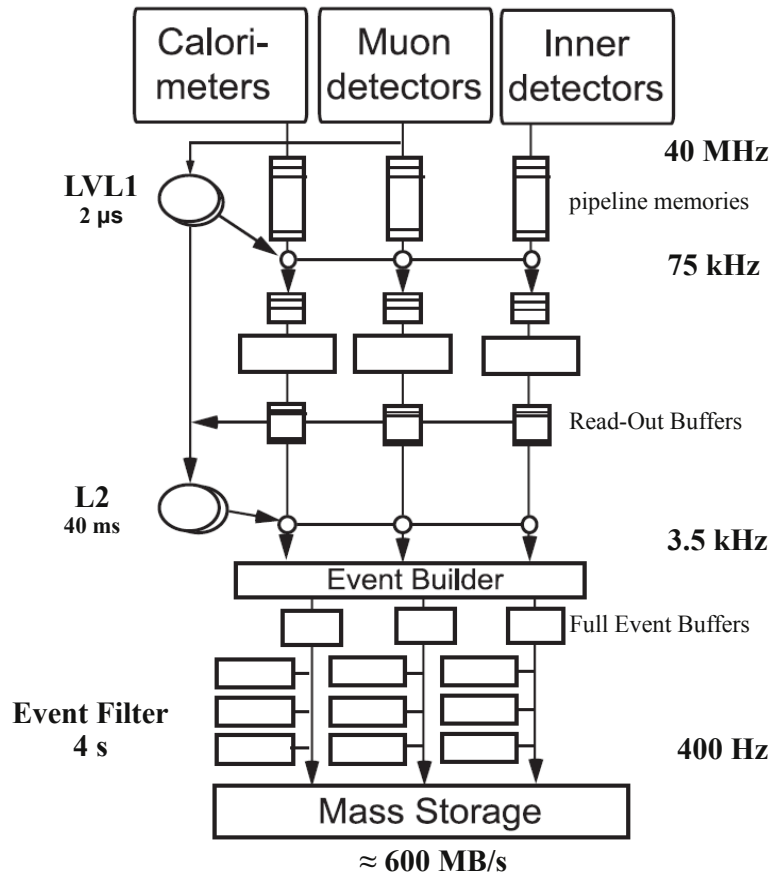


Figure 3.8: The ATLAS Trigger System [22].

Level-1 Trigger System

The Level-1 Trigger (LVL1) decision is hardware based and searches for interesting physics patterns such as high p_T electrons, muons, taus, jets or E_T^{miss} . The input to the LVL1 trigger is given by the Calorimeters and the Muon Spectrometer trigger system, via

Regions of Interest (ROI), i.e. detector regions where interesting physics signatures have been identified. Three low (from 6 GeV to 9 GeV) and three high (from 9 GeV to 35 GeV) transverse momentum thresholds are used in a coincidence logic implementation when a trigger decision is made using logical thresholds. LVL1 trigger provides the *luminosity block* number¹⁸ and reduces the event rate from 40 MHz to 75 kHz. The average LVL1 decision is made in 1 μ s (another 1 μ s delay is caused by the signal propagation via the optical cable network) [22].

Level-2 Trigger System

Using a software based implementation, the Level-2 Trigger (L2) performs a detailed re-analysis of the LVL1 decision via ROIs and confirms or rejects it. After the L2 decision has been made (≈ 40 ms) the event rate has been reduced to approximately 3.5 KHz [22].

Event Filter

The last trigger decision step belongs to the Event Filter. Based on a multitude of *trigger menu* configurations¹⁹ the event rate is lowered down to a few hundreds of Hertz within ≈ 4 seconds after the bunch crossing. At this level the full event information is used, with input from the whole detector (including the Inner Detector). Data events are sorted out via inclusive *streams* depending on trigger that fired the event²⁰. L2 together with the Event Filter are sometimes called **High Level Trigger** (HLT) [22].

3.9 ATLAS Forward Detectors

LUCID

Measuring relative luminosity via inelastic scattering and monitoring beam parameters, the LUCID detector²¹ is placed in the very forward region ($5.6 < |\eta| < 5.9$) where high-radiation levels are expected. LUCID is composed of two modules each situated at a distance of 17 m from the interaction point and contains a set of 20 aluminium polished tubes measuring 1.5 m in length filled with gaseous C₄F₁₀. The tubes are arranged symmetrically around the beam line and point towards the interaction point. The Cerenkov

¹⁸The *luminosity block* corresponds to a small period in time (few minutes) of data taking during which the overall detector conditions do not change significantly.

¹⁹A *trigger menu* is a selection criteria at trigger level.

²⁰The main streams used in ATLAS are: *electrons, gamma, muons, taus, jets* and E_T^{miss} .

²¹Luminosity measurement using Cerenkov Integrating Detector.

light signal produced by the charged particles passing through the active material is collected via photomultipliers placed at the end of each tube. With a small timing response per tube in the order of a few ns, LUCID can provide a good estimation of the number of interactions per bunch crossing²² - μ for values up to 30 [22, 28].

ZDC

The ZDC²³ was constructed for luminosity estimations in proton-proton and heavy ion collisions by measuring the neutral particle flux component (n, γ) emerging from the interaction point. It covers a $|\eta| > 8.3$ region and consists of two modules placed on each side of ATLAS at ± 140 m from the interaction point. Each module is structured into an electromagnetic calorimeter and three hadronic calorimeter layers [22, 28].

ALFA

Located at approximately 240 m from the nominal interaction point, the ALFA²⁴ detector provides absolute luminosity measurements via small angle elastic proton-proton scattering in the Coulomb-Nuclear Interference region.

The ALFA detector is structured in four Roman-pot spectrometers, two on each side of ATLAS with a separation of 4 meters, containing scintillating fibre trackers with spatial resolution of $30 \mu\text{m}$ [22].

3.10 The Worldwide LHC Computing Grid

The four major experiments at the LHC produce annually more than 20 Petabytes of data which will have to be stored and analyzed. This vast amount of data exceeds by far the storage possibilities of most computer clusters. Consequently, an international project - the **Worldwide LHC Computing Grid** (WLCG), has been developed as a high storage and computing capacity network which distributes the data in a hierarchical way down to the end-user. Today the WLCG has become a collaboration of 36 countries counting more than 170 computing farms. Data inside the WLCG structure is distributed via 4 levels, called *Tiers* as shown in figure 3.9. The first level - Tier-0, is essentially a high-storage computer-cluster located at CERN and has the role of recording and processing the raw data output of the Event Filter on tape. Tier-0 applies the first event reconstruction algorithm, calibrating and aligning the data before sending it further to the Tier-1. Distributed all around the world, Tier-1 is a network of computer farms which re-analyze and store

²²The bunch time gap in 2011 and 2012 data was 50 ns.

²³Zero Degree Calorimeter.

²⁴Absolute Luminosity For ATLAS.

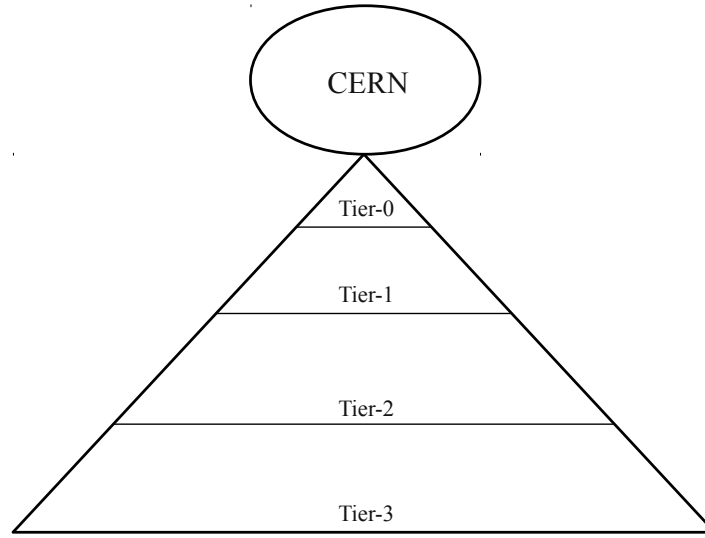


Figure 3.9: The Tier-structure of the Worldwide LHC Computing Grid.

large quantities of data. From here the data is sent at the Tier-2 level where only a fraction of it is stored. The Tier-2 is also responsible for MonteCarlo event simulation and user specific analysis. Finally, the Tier-3 consists of small computer clusters or even individual PCs dedicated to user analysis.

Driven by storage minimizing reasons, quite frequently user analysis jobs are sent to the Tier-2 grid infrastructure where the output is later picked up rather than replicating the data locally [29].

3.11 MonteCarlo Simulation

MonteCarlo data correspond to phenomenological-based simulation models of the real data produced at LHC and collected by the ATLAS detector. In a first phase, MonteCarlo events are created by dedicated event generators, simulating as close as it is reasonably possible real data events. At this step, the debris of proton-proton collisions such as initial and final state radiation, pileup, multiple parton interactions as well as the production of 4-vectors of particles originating from the hard scatter such as hadrons, electrons, gammas, muons, taus, takes place. Secondly, in order to account for detector effects originating in the interaction of particles with the detector material and the surrounding magnetic fields, all particles are passed through a GEANT4 simulation of the ATLAS detector.

The reconstruction of the ATLAS data as well as production of corresponding MonteCarlo events are done by the ATHENA software model using a Gaudi framework and

stored into raw object data (ROD) files. Driven by both disk storage as well as CPU time minimization, the ROD event file is further reduced since not all its available information is directly relevant for physics analysis. As a consequence the relevant ROD information is stored into smaller sized files. With a decreasing size there are event summary data (ESD), analysis object data (AOD) and derived physics data D3PD files.

Multipurpose MonteCarlo generators such as PYTHIA or HERWIG cover a wide spectrum of physical processes ranging from leading order matrix elements computation up to simulation models of the underlying event, parton shower and hadronization. Other generators have a narrower functionality, dealing only with matrix element estimations at leading-order such as for example ALPGEN [16], or at next-to-leading order - MC@NLO [17].

4 Data and MonteCarlo samples

4.1 Data Samples

This analysis uses 2011 proton-proton collision data at $\sqrt{s}=7$ TeV, collected with the ATLAS detector, corresponding to a recorded integrated luminosity of 5.25 fb^{-1} (see figure 4.1).

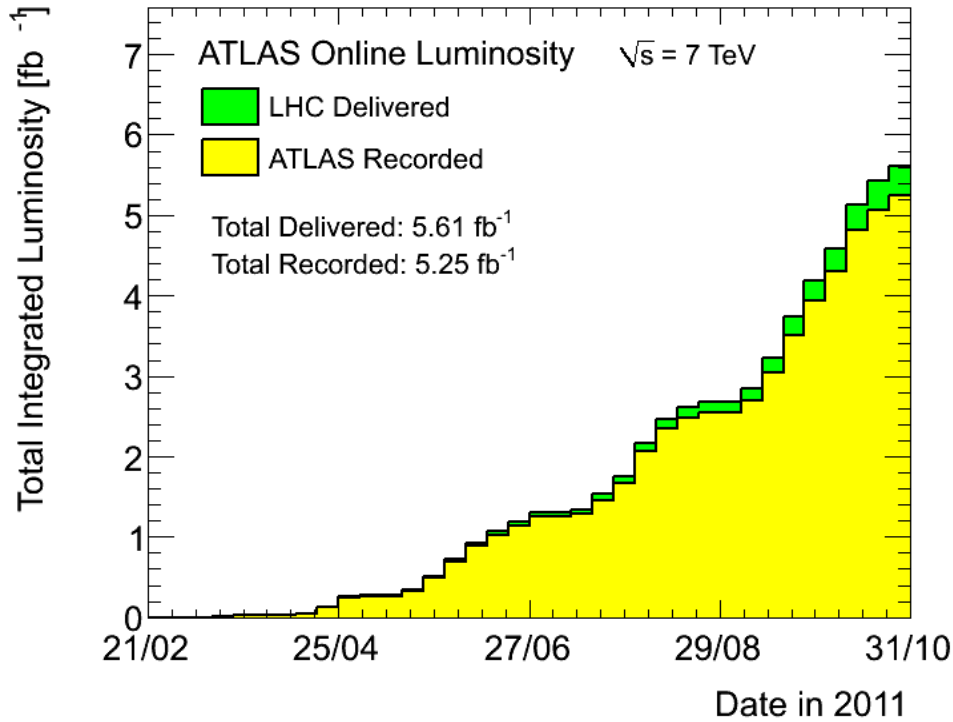


Figure 4.1: Integrated luminosity collected with the ATLAS detector during the year 2011 shown as a function of time. The instantaneous luminosity increase is driven by the higher bunch content while the bunch-spacing has remained at a constant 50 ns [12].

Driven by detector, trigger or beam conditions (such as the bunch charge or bunch multiplicity), the ATLAS data are splitted in a number of periods as shown in table 4.1. Between data periods E and H, six front-end boards (FEBs) of the LAr Calorimeter were nonfunctional, resulting in an inoperative detector region between $-0.1 < \eta < 1.5$ and $-0.9 < \phi < -0.5$. Starting with data period I, four out of the six faulty FEBs have been replaced. As a result of the missing LAr FEBs in data, the reconstruction and identification of physical-objects (especially electrons and jets) which happened to fall inside the affected calorimeter region became highly inefficient. A similar period dependent LAr faulty region (*LAr Hole*) was simulated in MonteCarlo. Events with jets with $p_T > 20$ GeV and $\Delta R_{jet-lepton} < 0.3$ (see section 5.1) falling inside the problematic LAr calorimeter region have been vetoed in both data and simulation since they can potentially lead to wrong missing transverse energy reconstruction.

The *Event Filter* output is divided and stored at CERN's Tier-0 in a number of physics-object data streams such as *electrons*, *gamma*, *muons*, *taus* or *jets* streams. The *electrons* and *gamma* streams, having similar features, have been merged into an inclusive *egamma* stream.

Since the $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$ decay channels of $Z + b\bar{b}$ events have been studied in this analysis exclusively, only the *egamma* and *muons* data streams have been used. The lowest unprescaled trigger configurations recommended by the ATLAS Standard Model W/Z working group [30] have been used to different data periods in which the pileup conditions changed significantly (see table 4.1). Di-electron and single-muon trigger configurations have been used in the electron and muon channels respectively ¹.

A fraction of approximately 10% of the recorded luminosity did not meet the Standard Model W/Z working group data quality conditions and has been consequently vetoed using a set of good runs lists (GRL).

Skimmed versions of the datasets ² provided by the Standard Model W/Z working group have been used.

4.2 MonteCarlo Samples

All MonteCarlo signal and background datasets considered in this analysis have been produced by the Standard Model W/Z working group during the third production campaign - MC11c, corresponding to the closest to-data simulation. Every generated event is passed

¹Di-muon trigger MonteCarlo scale-factors became available at a later stage of the analysis.

²A number of preselection cuts "skimming" (such as requiring at least two leptons of the same flavor, trigger and GRL selection) has been applied to both *egamma* and *muons* data streams.

Data stream	Data Period	Applied trigger	$\int L(t) dt$
<i>egamma</i>	D-J	EF_2e12_medium	1.659 fb ⁻¹
<i>egamma</i>	K	EF_2e12T_medium	0.583 fb ⁻¹
<i>egamma</i>	L-M	EF_2e12Tvh_medium	2.398 fb ⁻¹
<i>egamma</i>	D-M		4.640 fb ⁻¹
<i>muon</i>	D-I	EF_mu18_MG	1.435 fb ⁻¹
<i>muon</i>	J-M	EF_mu18_MG_medium	3.204 fb ⁻¹
<i>muon</i>	D-M		4.639 fb ⁻¹

Table 4.1: Period dependent triggers used in data and their corresponding integrated luminosity.

through a GEANT4 simulation of the ATLAS detector in order to simulate the interactions between particles and the detector material. Several MonteCarlo generator configurations simulating the matrix element (ME), parton shower (PS), hadronization and underlying event (UE) have been used for signal and background samples as shown in tables 4.2 and 4.3.

Zb \bar{b} -samples

Signal $Z + b\bar{b}$ samples (see Table 4.2) have been generated with ALPGEN using Leading Order matrix element (ME) computations for production of events with a pair of leptons³ in association with a pair of *massive* bottom-antibottom quarks ($m_{b(\bar{b})} = 4.7$ GeV). Additional *light*-partons can be produced in the ME reaching multiplicities up to three. A default mass of 0 GeV is chosen for all non-*b* quark flavors, including the charm. The parton shower is modeled with HERWIG while the Underlying Event is simulated with JIMMY. A matching between the matrix element and the partons shower is realized using the MLM method [33].

Two main processes contribute to the $Z + b\bar{b}$ production, namely gluon fusion (GF), and gluon splitting (GS) as shown in figure 2.7.

Z + light-parton samples

The most important background for this analysis comes from the production of a Z boson in association with light-flavor quarks. The $Z + \text{light-flavor}$ samples (see Table 4.3) (including also charm quarks) have been simulated with the same generator ALPGEN + HERWIG/ JIMMY configuration as for signal events. The generated light-quark multiplicity reaches values up to five and the invariant mass of the generated lepton falls inside a window of $40 \text{ GeV} < M_{ll} < 2000 \text{ GeV}$.

³ $30 \text{ GeV} < M_{ll} < 10000 \text{ GeV}$

Heavy Flavor Overlap Removal

Even though the $Z + \text{light partons}$ samples contain only light-flavor quarks at the ME level, a heavy-flavor component can be subsequently added to the event via GS in the HERWIG parton shower. An example of such processes is shown in figure 4.2, where ALPGEN generates a Z boson in association with a gluon which gets split into a pair of *massless* bottom-antibottom quarks by HERWIG. The same event is also to be found in Z+heavy-flavor signal samples (see figure 2.7 top-right plot), with the difference that here the entire process is generated exclusively by ALPGEN (ME) and the bottom-antibottom pair is *massive*. In order to avoid double-counting, common events to both sample sets have been rejected in one and kept in the other using the Heavy Flavor Overlap Removal (HFOR) tool decision [31]. Under the assumption that the Z+heavy-flavor event topology at large angular separation between the b -quarks is better described by the ALPGEN ME whilst at small $\Delta R_{b\bar{b}}$ the HERWIG parton shower GS yields superior event modeling, the HFOR decision is taken using ΔR as a discriminant:

- if $\Delta R_{b\bar{b}} \geq 0.4$ the $Z + b\bar{b}$ event is selected from ALPGEN Z+heavy-flavor samples.
- if $\Delta R_{b\bar{b}} < 0.4$ the $Z + b\bar{b}$ event is selected from ALPGEN Z+light-flavor samples.

As a result, a common $Z + b\bar{b}$ signal sample is formed with events originating from both Z+light-flavor and Z+heavy-flavor sample-sets.

“Other” samples

The second most important background component comes from top-antitop events which were simulated with the MC@NLO generator in association with HERWIG (parton shower and hadronization) and JIMMY (UE modeling). Only the di-leptonic and semi-leptonic top-antitop decays were selected.

Single-top as well as inclusive Wt events have also been considered and were simulated using the same MC@NLO + HERWIG/ JIMMY configuration as top-antitop samples. It has been shown that the t-channel single-top MC@NLO samples contained unphysical jet modeling in HERWIG parton shower. Consequently, a different generator configuration was used for these samples, namely ACERMC+PYTHIA using LO matrix element calculations. In both top-antitop and single-top samples the mass of the top-quark has a value of 172.5 GeV.

Diboson background processes such as WZ ($Z \rightarrow ll$, $W \rightarrow q\bar{q}$) and ZZ ($Z \rightarrow ll$, $Z \rightarrow q\bar{q}$) were also used, since they have a close-to-signal event topology. Diboson events were

simulated with MC@NLO generator (for ME) in association with HERWIG/JIMMY (PS, hadronization and UE).

All events have been scaled to NLO cross-sections using k-factors ⁴ provided by the ATLAS TopWorkingGroup [32].

Sample ID	Physical Process	Event Generator(s)	σ [pb]	k-factor	N_{evts}
109300	$Zb\bar{b}+0$ part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	6.57	1.25	409999
109301	$Zb\bar{b}+1$ part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	2.48	1.25	160000
109302	$Zb\bar{b}+2$ part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	0.89	1.25	60000
109303	$Zb\bar{b}+3$ part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	0.39	1.25	30000
109305	$Zb\bar{b}+0$ part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	6.56	1.25	409949
109306	$Zb\bar{b}+1$ part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	2.47	1.25	155000
109307	$Zb\bar{b}+2$ part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	0.89	1.25	60000
109308	$Zb\bar{b}+3$ part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	0.39	1.25	29999

Table 4.2: Signal samples identification number, underlying physical process, Event Generator, lepton filter efficiency times cross-section, k-factor and the number of generated events.

Every simulation sample in the MC11c production campaign has been generated with a pileup event-content close to that of the 2011 data. It has been shown that the simulation event pileup content has slightly overestimated that in data. To correct for the observed discrepancy, ad-hoc corrections have been applied to all events. Figure 4.3 shows the pileup event distribution in data and simulation before and after the corresponding corrections.

In order to closely follow the data event selection, the same period-dependent trigger configuration as in data has been applied to an equivalent MonteCarlo event fraction.

⁴The k-factor is defined as the ratio between the NLO and LO cross-sections for a certain physical process.

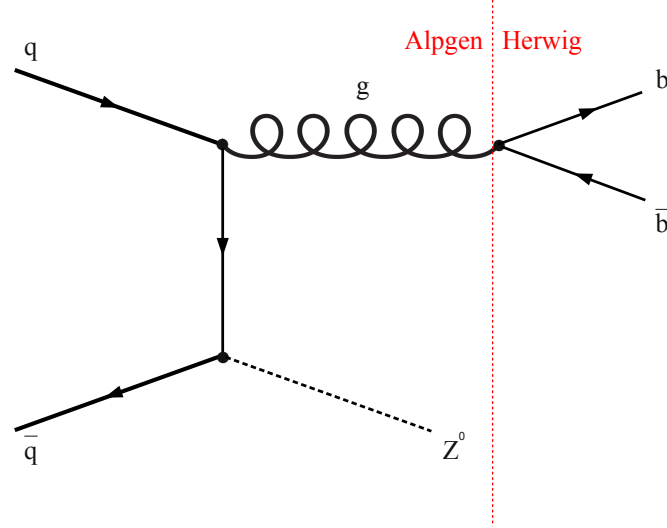


Figure 4.2: A $Z + b\bar{b}$ event is produced from an initial Z +gluon Alpgen event via gluon-splitting in the parton shower.

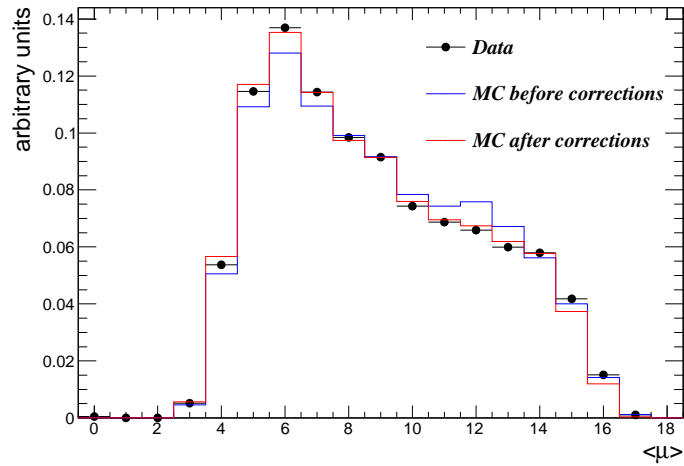


Figure 4.3: Average number of interactions per bunch crossing - $\langle\mu\rangle$, in data (dots) and simulation, before (blue curve) and after (red curve) corrections.

Sample ID	Physical Process	Event Generator(s)	$\epsilon_F \cdot \sigma$ [pb]	k-factor	N_{evts}
107650	Z+0 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	668.32	1.25	6618284
107651	Z+1 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	134.36	1.25	1334897
107652	Z+2 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	40.54	1.25	2004195
107653	Z+3 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	11.16	1.25	549949
107654	Z+4 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	2.88	1.25	149948
107655	Z+5 part., $Z \rightarrow e^+e^-$	ALPGEN+JIMMY	0.83	1.25	50000
107660	Z+0 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	668.68	1.25	6615230
107661	Z+1 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	134.14	1.25	1334296
107662	Z+2 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	40.33	1.25	1999941
107663	Z+3 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	11.19	1.25	309899
107664	Z+4 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	2.75	1.25	35000
107665	Z+5 part., $Z \rightarrow \mu^+\mu^-$	ALPGEN+JIMMY	0.77	1.25	50000
105200	$t\bar{t}$	MC@NLO+JIMMY	79.01	1.146	14973835
117360	Single-top, ev t-channel	ACERMC+PYTHIA	8.06	0.865	999295
117361	Single-top, $\mu\nu$ t-channel	ACERMC+PYTHIA	8.06	0.865	999948
108343	Single-top, ev s-channel	MC@NLO+JIMMY	0.47	1.064	299948
108344	Single-top, $\mu\nu$ s-channel	MC@NLO+JIMMY	0.47	1.064	299998
108346	Wt inclusive	MC@NLO+JIMMY	14.59	1.079	899694
105930	$ZZ, \ell\ell qq$	MC@NLO+JIMMY	0.559	-	25000
105942	$W^+Z, qq\ell\ell$	MC@NLO+JIMMY	0.5415	-	24950
105972	$W^-Z, qq\ell\ell$	MC@NLO+JIMMY	0.2944	-	100000

Table 4.3: Background samples identification number, underlying physical process, Event Generator, lepton filter efficiency times cross-section, k-factor and the number of generated events.

5 Object Reconstruction and Event Selection

This chapter is structured in two sections. In the first section the reconstruction and identification of physical objects used in the analysis is presented together with a description of the selection cuts and the object corrections applied to simulation. The second section will focus on the signal event reconstruction, some further event corrections and finally at the end of the section a number of plots is shown for the signal event selection at different steps of the analysis.

5.1 Reconstruction and Event Preselection

Primary Vertex

Reconstruction and Identification

The presence of multiple vertices in the same bunch crossing due to pileup effects requires a clear definition for the hard scatter point, i.e. the Primary Vertex (PV). The reconstruction of the PV is based exclusively on tracking information inside the Inner Detector and, simplistically, the PV is defined as the vertex with the highest sum of p_T^2 of the tracks, where each track must pass a $p_T > 100$ MeV threshold. Further track constraints such as on the number of Pixel, SCT or TRT hits, are required. Once the event's PV is found, the angular coordinates η and ϕ of physical objects can be defined with respect to it [34].

Selection

In order to reject primary vertices induced by cosmic rays or beam halo effects, a minimum number of at least three good-quality tracks each with $p_T > 400$ MeV is required for the primary vertex reconstruction in this analysis.

Electrons

Reconstruction and Identification

The electron reconstruction is based on a set of off-line algorithms seeded by energy deposits in the EM calorimeter system and/or tracks inside the Inner Detector. Cluster based electron candidate reconstruction algorithms such as *Author 1* or *Author 3* [35], start with electromagnetic *tower*¹ seeds using a $3 \times 7 - \Delta\eta \times \Delta\phi$ sliding window algorithm based on energy deposits inside the middle EM calorimeter layer until a local maximum is reached. The *Author 3* electron reconstruction algorithm requires the electron cluster to be matched with an ID track (if track $p_T > 5$ GeV), *Author 2* is based exclusively on tracking information [35] and *Author 1* uses cluster energy deposits. Tracks compatible with photon-pair conversion inside the Inner Detector material are rejected by requiring b-layer Pixel hits [36]. The energy loss caused by interactions with the detector material varies as a function of η . Particularly the electron candidates with $|\eta_{cluster}| < 2.47$ excluding the transition region between the barrel and the end-caps ($1.37 < |\eta|_{cluster} < 1.52$) are reconstructed with a worse relative energy resolution.

The identification of electron candidates is split in three inclusive classes - *loose*, *medium* and *tight*² by using a set of E_T and η dependent track and shower parameters such as the electromagnetic *shower shape* and the *shower width*, the *hadronic leakage* (ratio between the E_T as measured in the Hadronic and EM calorimeters), the ratio between the p_T measured in the Inner Detector and the EM calorimeter E_T or the number of semiconductor (Pixel + SCT) and TRT hits [36, 37]. Additional requirements have been applied to electrons in 2011 data due to the increased event pileup content, resulting in a new set of optimized classes such as *loose++*, *medium++* and *tight++* [37]. Having the highest efficiency, *loose++* electrons also yield the smallest “fake”³ rejection rate. Using more stringent cluster and track requirements, *tight++* electrons show the highest fake rejection rate but, consequently, yield a smaller efficiency. As a compromise between a good overall efficiency and high-fake rejection, *medium++* electron candidates have been chosen for this analysis.

Electron reconstruction and identification efficiencies in simulation have been shown to

¹ ATLAS calorimeters are segmented in a 2-D *tower* grid network, each *tower* containing an η dependent number of calorimeter cells. Typical electromagnetic towers have a $0.025 \times 0.025 - \Delta\eta \times \Delta\phi$ granularity while hadronic towers have a coarser 0.1×0.1 segmentation.

² Electrons tagged as *tight* are a subclass of *medium* which in turn are a sub-class of *loose* electrons.

³ A jet’s energy deposition inside the EM Calorimeter can be misidentified with the electromagnetic energy cluster induced by an electron with a non-zero probability.

overestimate those in data, therefore ad-hoc event scale factors are applied to simulated events in order to minimise the observed differences [38].

Selection

Only electron tracks with *longitudinal impact parameter*⁴ values up to 1 mm and *transverse impact parameter* significance⁵ $|\sigma(d_0)| < 10$ are selected in order to increase the Z boson origin probability of the electron as well as reducing the multi-jets background component. The transverse energy of the electron is reconstructed as follows:

- if the number of Pixel + TRT hits ≥ 4 , the electron 4-vector is reconstructed using the calorimeter energy and η and ϕ coordinates from the Inner Detector track [39].
- otherwise, the electron 4-vector is reconstructed exclusively using calorimeter input [39].

Electron candidates reconstructed with *Author* 1 or 3 are required to pass a transverse energy threshold of 20 GeV within a pseudorapidity of $|\eta|_{cluster} < 2.47$ (excluding the barrel-end-cap region $1.37 < |\eta|_{cluster} < 1.52$).

A p_T and η dependent correction has been applied to the electron four-momenta in data as a systematic shift. Subsequent to the energy scale corrections, the electron four-momentum components show a worse resolution⁶ compared to those in simulation. Consequently, momentum corrections or “*smearing*” for electron objects in simulation have been applied following the recommendations given by the ATLAS ElectronGamma Physics Group [40]. Figure 5.1 top, shows the invariant mass distribution of the leading (by p_T) two electrons in simulation before (blue dotted line) and after (continuous red line) applying the smearing corrections compared to the distribution in data (dots). An improvement in the invariant mass distribution can be observed after applying the smearing correction.

To account for the 5%-10 % energy reconstruction underestimation in both data and simulation an additional correction for electrons in the transition region, $1.42 < |\eta|_{cluster} < 1.55$ has been applied [40].

⁴The impact parameter of a track with respect to the PV is defined as the distance at the perigee between the track extrapolation and the PV. The impact parameter measured in the plane transverse to the beam line is called *transverse impact parameter* while the one measured along the beam line is called *longitudinal impact parameter*.

⁵Computed as the *transverse impact parameter* divided by its uncertainty.

⁶Caused primarily by an underestimation of the constant term of the EM Calorimeter in simulation.

Muons

Reconstruction and Identification

Three different classes are used in muon reconstruction and identification, namely *stand-alone*, *combined* and *segment tagged* muons [42].

The *Stand-alone* muon class is seeded by track segments reconstructed from Muon Spectrometer stations which are subsequently extrapolated to the primary vertex, taking into account energy losses in the calorimeters and multiple scattering effects within the detector material.

Segment Tagged muon reconstruction is seeded by ID tracks extrapolated to the Muon Spectrometer or calorimeter system where they are matched using the closest muon track or energy deposit. *Segment Tagged* algorithms have been shown to yield superior reconstruction efficiencies, especially in regions where the Muon Spectrometer has uninstrumented areas such as at $|\eta| \sim 0$ and $|\eta| \sim 1.2$ [42].

The *Combined* muon class is seeded by Muon Spectrometer as well as Inner Detector tracks. In a first step tracks reconstructed inside the Muon Spectrometer from chamber segments are extrapolated to the interaction point where they are matched with Inner Detector tracks ($|\eta|_{\text{track}} < 2.5$) into combined tracks. Track matching is done by statistically combining track 4-vectors using a chi-squared fit. A second set of reconstruction algorithms is seeded by ID tracks extrapolated to the MS, improving the combined track fit. One of the largest used algorithm chains providing *combined* muon reconstruction in ATLAS is the STACO chain, yielding average muon reconstruction efficiencies of $\sim 98\% - 99\%$ [42]. Having the highest purity, STACO muons have also been the choice for this analysis.

Selection

Only STACO muons with $p_T > 20$ GeV and $|\eta| < 2.4$ are selected requiring the Inner Detector track to have a longitudinal impact parameter $|\Delta z| < 1$ mm and transverse impact parameter significance $|\sigma(d_0)| < 3$. The last two cuts have the role of increasing the Z boson origin probability of the muon, reducing multi-jets background events and veto cosmic muons. The rejection of muons originating from semileptonic decays from heavy quarks is realized via an Inner Detector-based isolation variable requiring that the ratio between the sum of p_T of all tracks around the muon in a cone with radius $\Delta R = 0.2$ ($p_T^{\text{cone}20}$), and the muon p_T be less than 10%. An additional set of Inner Detector track criteria such as on the number of b-layer Pixel hits, or SCT + TRT hits is required in order to reject muons with beam or cosmic origin [43].

Similarly to electrons, reconstruction and identification efficiency scale factors as well as momentum resolution smearing in simulation have been applied in order to improve

the agreement with data. Figure 5.1 bottom, shows the invariant mass distribution of the leading (by p_T) two muons in simulation before (blue dotted line) and after (continuous red line) applying the smearing corrections compared to the distribution in data (dots). A clear improvement can be observed. The analysis correction packages have been provided by the ATLAS Muon Combined Performance Group [43].

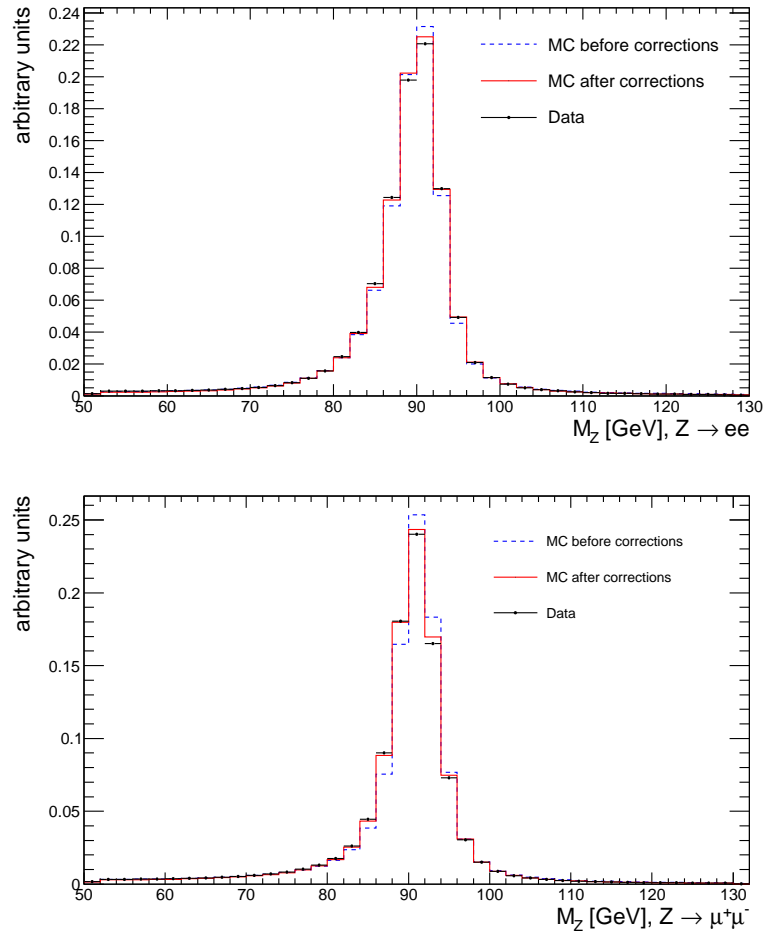


Figure 5.1: Invariant mass distribution for the leading two electrons (top figure) and leading two muons (bottom figure). The blue dotted lines (continuous red line) correspond to distributions prior (subsequent) to the smearing corrections in simulation. The black full dots correspond to the invariant mass distribution in data. All distributions have been normalized to an area of one event.

Jets

Reconstruction and Identification

Jets have been reconstructed with the anti- k_T [44] algorithm from topological calorimeter clusters ⁷ [45] using a distance parameter of $R = 0.4$. The anti- k_T algorithm starts by reconstructing all momentum-space distances between calorimeter objects i and j as well as the distance between object i and the beam (B):

$$d_{ij} = \min\left(p_{Ti}^{-2}, p_{Tj}^{-2}\right) \frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = p_{Ti}^{-2}, \quad (5.1)$$

with $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$. If d_{ij} is smaller than d_{iB} , objects i and j are merged into an inclusive jet, otherwise i is called a jet and removed. The iterative process continues until all objects have either been clustered within jets or become jets on their own. The anti- k_T algorithm has been shown to be infrared and collinear safe as well as soft-radiation resilient at jet boundaries [44].

Subsequent to the *electromagnetic scale (EM Scale)* calibration ⁸, an additional hadronic **Jet Energy Scale (JES)** is applied to all reconstructed jets in the event. In order to correct for the residual differences between the combined **EM+JES** scale jet calibration in simulation and data, an additional calibration is applied to jets with $p_T^{jet} > 20$ GeV and $|\eta^{jet}| < 4.5$ using offset and beam-spot corrections as recommended by the ATLAS Jet and Missing Et Group [46], resulting in a modification of the original jet's 4-momentum components.

Selection

In order to minimise pileup contamination effects, all selected jets in this analysis have been required to have at least 75% of their τ track sum originating from the Primary Vertex using the “*jet vertex fraction*” (JVF) variable.

Events with jets with $p_T > 20$ GeV pointing to known faulty calorimeter regions such as in the *LAr Hole* have been rejected.

⁷Topological clusters correspond to a 3-dimensional array of neighboring calorimeter cells with signal content significantly above the noise threshold. The cell-number associated to a topo-cluster unlike that of towers is a dynamic parameter.

⁸In a first phase all jet-objects in ATLAS are calibrated at the *electromagnetic scale (EM Scale)*, i.e. the energy scale corresponding to electromagnetic showers, ignoring differences in response between EM and hadronic objects or the eventual energy losses from non-instrumented or dead detector regions.

B-Tagging

Reconstruction and Identification

Selecting bottom-quark induced jets as well as rejecting those with different quark origins is realized with the help of *b*-tagging algorithms. *B*-tagging algorithms exploit the properties of *b*-hadrons such as their specific long life-time⁹ $\sim 10^{-12}$ s before decaying into lighter hadrons, producing thus a secondary vertex (SV) which can be reconstructed from the decay tracks inside the Inner Detector with high precision (see Figure 5.2).

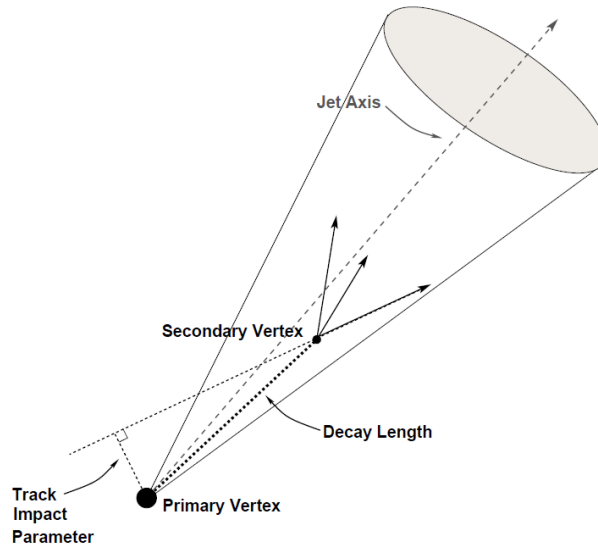


Figure 5.2: Secondary vertex decay topology [51]. Tracks originating from the SV decay will tend to have larger values of the impact parameter significance compared to those emerging from the PV decay.

Additional characteristics which seed *b*-tagging algorithms are the high track multiplicity of the SV, the secondary vertex mass or the track's longitudinal and transverse *impact parameter* with respect to the PV. A wide spectrum of algorithms has been developed, starting with relatively simple taggers such as IP3D+SV1 [48], SV0 [49] or JetFitter [48] up to more complex taggers such as JetFitterCOMBNN [48] or MV1 [50].

⁹The relative long life-time of weakly decaying *b*-hadrons is a direct consequence of the CKM matrix element V_{cb} suppression.

JetFitterCOMBNN

JetFitterCOMBNN is a neural network algorithm trained to distinguish between light, charm and bottom induced jets. The output of the neural network is interpreted in terms of the probability for a certain jet to have been induced by a light (pu), charm (pc) or bottom (pb) quark. The tagger's response is seeded by the outputs provided by two simpler b -tagging algorithms, namely JetFitter¹⁰ and IP3D¹¹.

MV1

The decision of the MV1 b -tagger is based on a multivariate analysis seeded by the outputs of the IP3D, SV1 and JetFitterCOMBNN algorithms. The MV1 discriminant combines the input parameters and generates a continuous distribution which tends to values close to 0 for light flavored jets or 1 for b -induced jets.

Selection

Having shown the highest performance, the MV1 tagger has been chosen for this analysis at an 75% efficiency operating point¹².

In order to correct both the b -tagging efficiency (tagged jets) and the b -tagging inefficiency (the percentage of jets which are not tagged) to those in data, flavor-dependent correction scale factors in bins of p_T and η have been applied in simulation to all selected jets in the event following the ATLAS Flavour Tagging Working Group recommendations [52]. Since the b -tagging efficiency/inefficiency corrections are provided for jets down to 20 GeV, a $p_T > 20$ GeV threshold has been applied to all selected b -tagged jets calibrated at the hadronic **JES**. A rapidity cut $|y| < 2.4$ has been applied to all b -tagged jets in order to reach the Inner Detector reconstruction efficiency plateau.

The highest reducible background for this analysis comes from the production of a Z boson in association with *light*-flavor¹³ or c -induced jets. Therefore the use of b -tagging plays a crucial role in the reduction of the Z + *light*-jets and Z + c -jets backgrounds.

¹⁰JetFitter algorithm uses the assumption that the decay vertex of the b -hadron, the c -hadron and the PV lie approximately on the same line. The output discriminant is based on a number of parameters, such as the invariant mass of the SV, the number of secondary vertices reconstructed from at least two tracks, the number of tracks used to reconstruct the SV and the energy ratio of tracks used in the SV reconstruction to the total number of tracks inside the jet.

¹¹The IP3D decision uses distributions of the longitudinal and transversal impact parameter significances of tracks inside a jet for the reconstruction of the secondary vertex.

¹²The 75% value of the operating point has been derived using a top-antitop simulation sample. The b -tagging efficiencies do not necessarily yield the same values for other processes such as Z + $b\bar{b}$.

¹³*Light* flavored jets originate from u , d , s -quarks or *gluons*.

Missing Transverse Energy

Reconstruction

The energy imbalance in the plane transverse to the beam direction, reconstructed from calorimeter cells, MS and ID tracks is called *missing transverse energy* - E_T^{miss} :

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad \text{with :} \quad (5.2)$$

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,muon} \quad \text{and :} \quad E_{x(y)}^{miss} = \sum_{\text{all objects}} -E_{x(y)}. \quad (5.3)$$

where E_x^{miss} and E_y^{miss} are the x and y components of the E_T^{miss} term. The calorimeter component - $E_{x(y)}^{miss,calo}$ is built from all topological cluster calorimeter cells ($|\eta| < 4.5$) calibrated at the electromagnetic energy scale (electrons, gammas) or the hadronic scale (non-leptonic decaying taus or hadrons) taking into account also the eventual muon energy depositions. The muon component - $E_{x(y)}^{miss,muon}$ is seeded by all muon-objects from both the Muon Spectrometer ($|\eta| < 2.7$) and the Inner Detector. *Combined* muons are used for $|\eta| < 2.5$ (ID acceptance), while outside this region, $2.5 < |\eta| < 2.7$, they are replaced with *stand-alone* muons. In order to avoid double-counting the calorimeter energy deposit was excluded from the muon $E_T^{miss,calo}$ term [53].

Selection

Having a large cross-section and a similar to signal final state, top-antitop events¹⁴ constitute an important and irreducible background in this analysis. Events with $E_T^{miss} > 30$ GeV have been rejected in order to reduce the $t\bar{t}$ background component¹⁵.

Since a number of physical objects used in the analysis requires corrections which modify the original 4-momentum components, such as smearing or energy scale corrections, the E_T^{miss} will have to be redefined taking these corrections into account. This is realized by propagating all the 4-vectors affected by such corrections to a *MissingETUtility* tool provided by the ATLAS Standard Model Working Group [55] which returns the corrected E_T^{miss} value.

¹⁴The inclusive $t\bar{t}$ cross-section at $\sqrt{s} = 7$ TeV with the ATLAS detector has been measured to be $\sigma_{t\bar{t}} = 145 \text{ pb} \pm 31 \text{ pb}(\text{stat.})^{+42}_{-27} \text{ pb}(\text{syst.})$ [54].

¹⁵Due to the presence of final-state neutrinos, semileptonic and dileptonic $t\bar{t}$ channels are usually characterised by a large missing transverse energy signature.

5.2 Final Event Selection

Z Boson Reconstruction

Events with exactly two opposite charged same-flavor leptons passing the selection criteria described in the previous section are used for the Z boson candidate reconstruction. The mass of the Z boson candidate must lay inside a 76 GeV - 106 GeV window. The event is vetoed if additional isolated leptons (of any flavor) passing the selection cuts are found in the event.

B-Jets

At least two b -tagged jets selected via the MV1 algorithm passing all the jet selection cuts are required in each event.

Further Event Corrections

Events with bad-quality jets [47] with $p_T > 20$ GeV and $\Delta R_{jet,signal-lepton} < 0.3$ have been rejected in both data and simulation since such events have been shown to degrade the E_T^{miss} resolution [57]. This procedure will be referred to in the following as E_T^{miss} *cleaning*. Additionally selected jets within $\Delta R < 0.5$ to a selected lepton have been rejected as the presence of a close-by prompt lepton can deteriorate the jet 4-momentum reconstruction.

Trigger efficiency scale factors ¹⁶ have been applied to every simulation event which fired the trigger. For the p_T and η lepton thresholds used in this analysis which fall inside the trigger efficiency plateau, the overall scale factors are within 1% of unity. In addition, on both data and simulation a trigger matching requirement was imposed.

EvtGen Reweighting

It has been shown [66] that the decay of b -hadrons is mismodeled in samples in which Herwig was used to simulate the parton shower and hadronization processes. As a result, the distribution of the charged particle multiplicity originating from weak b -hadron decays is noticeably underestimated in simulation with respect to data. In order to correct for the observed mismodeling, event-by-event ad-hoc scale factors have been applied in samples in which Herwig was used to model the parton shower and hadronization processes,

¹⁶Showing strong p_T and η dependency, trigger scale factors have been derived by the ATLAS Trigger and Data Acquisition [56] using tag-and-prove comparison methods between data and simulation.

namely top-antitop and $Z + b\bar{b}$ events. The correction scale factors have been obtained from an equivalent set of signal and top-antitop samples as the default ones (shown in tables 4.2 and 4.3) in which the weak b -hadron decays have been modeled with the EvtGen package¹⁷ [65]. Such corrections are necessary since the b -tagging algorithms used in this analysis are based on the input on the distribution of charged particle multiplicities.

Results

Table 5.1 gives a summary of the selection cuts applied in the electron and muon channels.

Next, tables 5.2 and 5.3 give the electron and muon channel event yields for data and simulation at different selection levels in the analysis. The third (MC Non-scaled) and fourth (MC Scaled) columns represent the total simulation event yields prior and subsequent to using the measured $Zb\bar{b}$ correction (for more details consult section 7.1) while the fifth column (Observed) corresponds to the data yields. A scale factor of 1.038 was applied in the electron channel to signal events in which at least two jets were matched with b -hadrons¹⁸ and a scale factor of 1.027 was applied to signal events with less than two jets matched with b -hadrons. The scale factors have been estimated as the ratio between the post-fit and the pre-fit templates yield (see Section 7.1 for more details). For the muon channel the corresponding event scale factors were estimated to 1.30 (signal events with at least two jets matched with b -hadrons) and 1.048 for the remaining signal events. The second column, $Zb\bar{b}$, gives the scaled signal event yield contribution emerging from the $Zb\bar{b}$ samples (see table 4.2).

The third and fourth columns from tables 5.4 and 5.5 show the individual signal and background sample contributions to an event selection with a Z boson candidate and at least two b -tagged jets, before “MC non-scaled” and after “MC scaled”¹⁹ using the measured cross-section. The fifth column gives the observed event yield.

Figures 5.3 to 5.15 show various object kinematic distributions in the electron and muon channels for data and simulation in the final event selection region unless stated otherwise. All distributions have been weighted with the corresponding measured cross-section in both channels. A good description of the data by simulation is observed after using the $Z + b\bar{b}$ measured cross-section. In *Appendix B* a detailed study of the data and simulation

¹⁷The EvtGen package is expected to have a better modeling for the distribution of the charged particle multiplicity from b -hadron decays than Herwig.

¹⁸Each b -tagged jet was required to be matched with a weakly decaying b -hadron with $p_T^{b\text{-hadron}} > 5$ GeV, if the b -hadron was found within $\Delta R_{b\text{-hadron}, \text{jet}} < 0.3$ to the jet.

¹⁹Corresponding to the measured signal event yields.

comparison before using the measured cross-section in simulation is presented.

	Electron Channel	Muon Channel
Signal	EF_2e12_medium	EF_mu18_MG
	EF_2e12T_medium	EF_mu18_MG_medium
	EF_2e12Tvh_medium	Combined muons
	Author 1 or 3	Author 6
Leptons	$ \Delta z_0 < 1\text{mm}, \sigma d_0 < 10$	$ \Delta z_0 < 1\text{mm}, \sigma d_0 < 3$
	medium++	$p_T^{\text{cone20}}/p_T < 10\%$
	—	ID quality track selection
	$ \eta_{\text{cluster}} < 2.47$	$ \eta < 2.4$
Selection	veto $1.37 < \eta_{\text{cluster}} < 1.52$	—
	$p_T > 20\text{ GeV}$	$p_T > 20\text{ GeV}$
	exactly 2 electrons	exactly 2 muons
	veto additional leptons;	veto additional leptons
	$q_{e1} \cdot q_{e2} < 0$	$q_{\mu1} \cdot q_{\mu2} < 0$
	$76\text{ GeV} < M_{e^+e^-} < 106\text{ GeV}$	$76\text{ GeV} < M_{\mu^+\mu^-} < 106\text{ GeV}$
Jet	JVF > 0.75	
	LAr hole veto	
	$p_T > 20\text{ GeV}$	
	$ \eta < 2.4$	
Selection	0.75% MV1 tagger weight cut	
	$\Delta R_{\text{jet},\text{signal-lepton}} > 0.5$	
	$\geq 2\text{ }b\text{-jets}$	
E_T^{miss} selection	$E_T^{\text{miss}} < 30\text{ GeV}$	$E_T^{\text{miss}} < 30\text{ GeV}$
		$E_T^{\text{miss}} < 30\text{ GeV}$

Table 5.1: Selection cuts summary.

Selection	$Zb\bar{b}$	MC non-scaled	MC Scaled	Data
Total events	60064.9	5507581.6	5640685.3	1271155
GRL	60064.9	5507581.6	5640685.3	1271155
Primary Vertex	57599.1	5453035.5	5586139.2	1266840
LAr Error	57599.1	5453035.5	5586139.2	1263293
Trigger	30039.1	1954972.2	2007414.5	1263293
Z boson	16476.8	1077921.7	1106977.4	1126957
E_T^{miss} cleaning	16476.8	1077921.7	1106977.4	1125773
LAr Hole veto	16476.8	1077921.7	1106977.4	1124494
$E_T^{miss} < 30$ GeV	14192.7	965405.7	991447.5	1008769
$Z + \geq 1$ jet	7163.2	214510.6	220281.1	224124
$Z + \geq 1$ b -jet	3638.8	14642.6	15034.9	16723
$Z + \geq 2$ b -jets	401.2	670.9	691.4	692

Table 5.2: Electron channel cut-flow table for simulation before/after (MC non-scaled/MC Scaled) using the measured signal cross-section. The second column ($Zb\bar{b}$) shows the scaled signal event yields while the fifth column corresponds to the data.

Selection	$Zb\bar{b}$	MC non-scaled	MC Scaled	Data
Total events	61046.2	5481507.1	5719636.8	1831592
GRL	61046.2	5481507.1	5719636.8	1831592
Primary Vertex	58603.1	5416767.5	5654897.2	1823782
LAr Error	58603.1	5416767.5	5654897.2	1818811
Trigger	46315.5	3598584.9	3766397.3	1818811
Z boson	23218.4	1522052.3	1596038.1	1569827
E_T^{miss} cleaning	23218.4	1522052.3	1596038.1	1568250
LAr Hole veto	23218.4	1522052.3	1596038.1	1566523
$E_T^{miss} < 30$ GeV	19787.4	1354325.5	1420187.4	1395551
$Z + \geq 1$ jet	9922.6	297226.5	311720.4	305210
$Z + \geq 1$ b -jet	5066.3	19590.9	20636.5	22329
$Z + \geq 2$ b -jets	624.8	863.7	1011	1011

Table 5.3: Muon channel cut-flow table for simulation before/after (MC non-scaled/MC Scaled) using the measured signal cross-section. The second column ($Zb\bar{b}$) shows the scaled signal event yields while the fifth column corresponds to the and data.

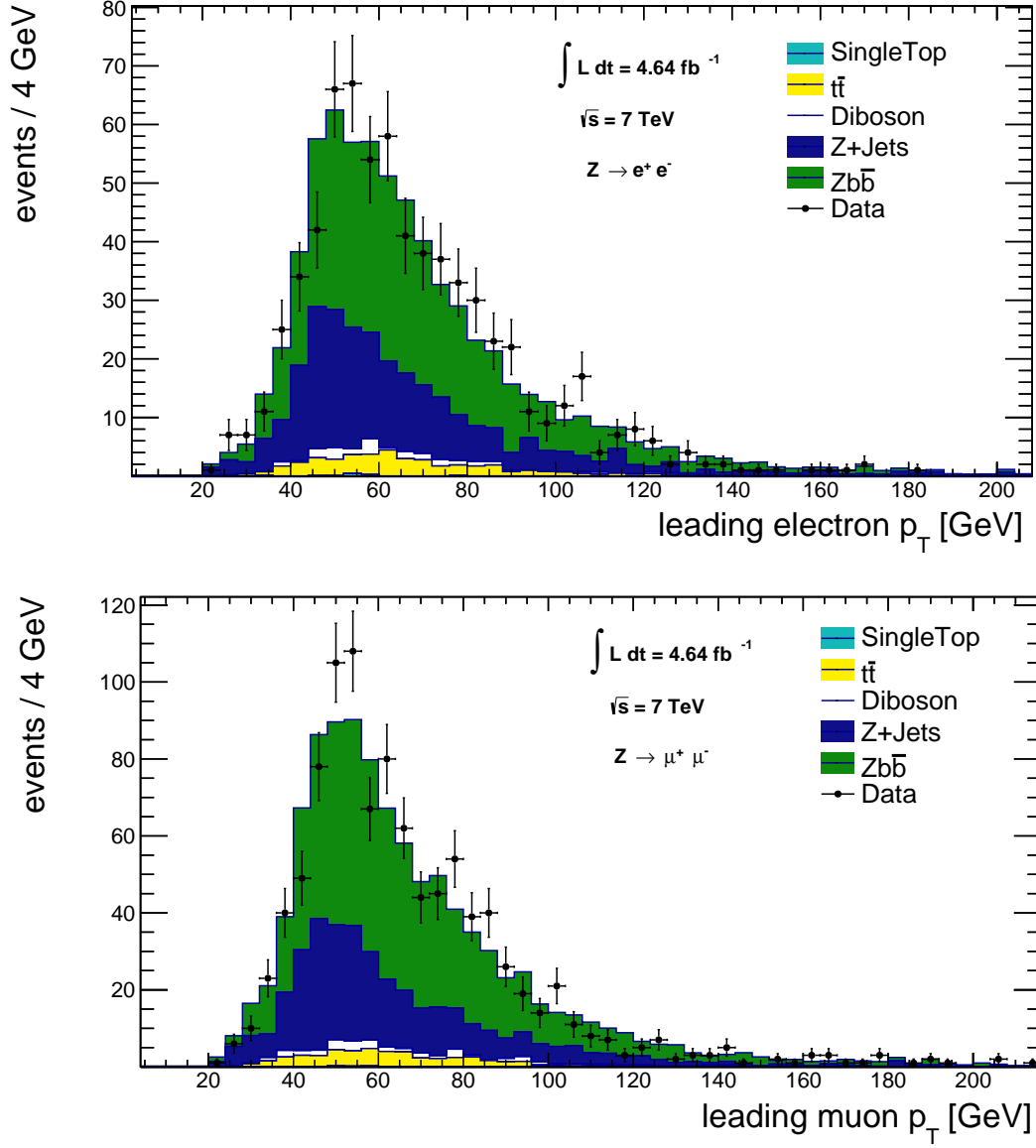


Figure 5.3: Data (points) versus simulation (stacked histograms) distributions of the leading lepton p_T in the electron channel (top plot) and the muon channel (bottom plot).

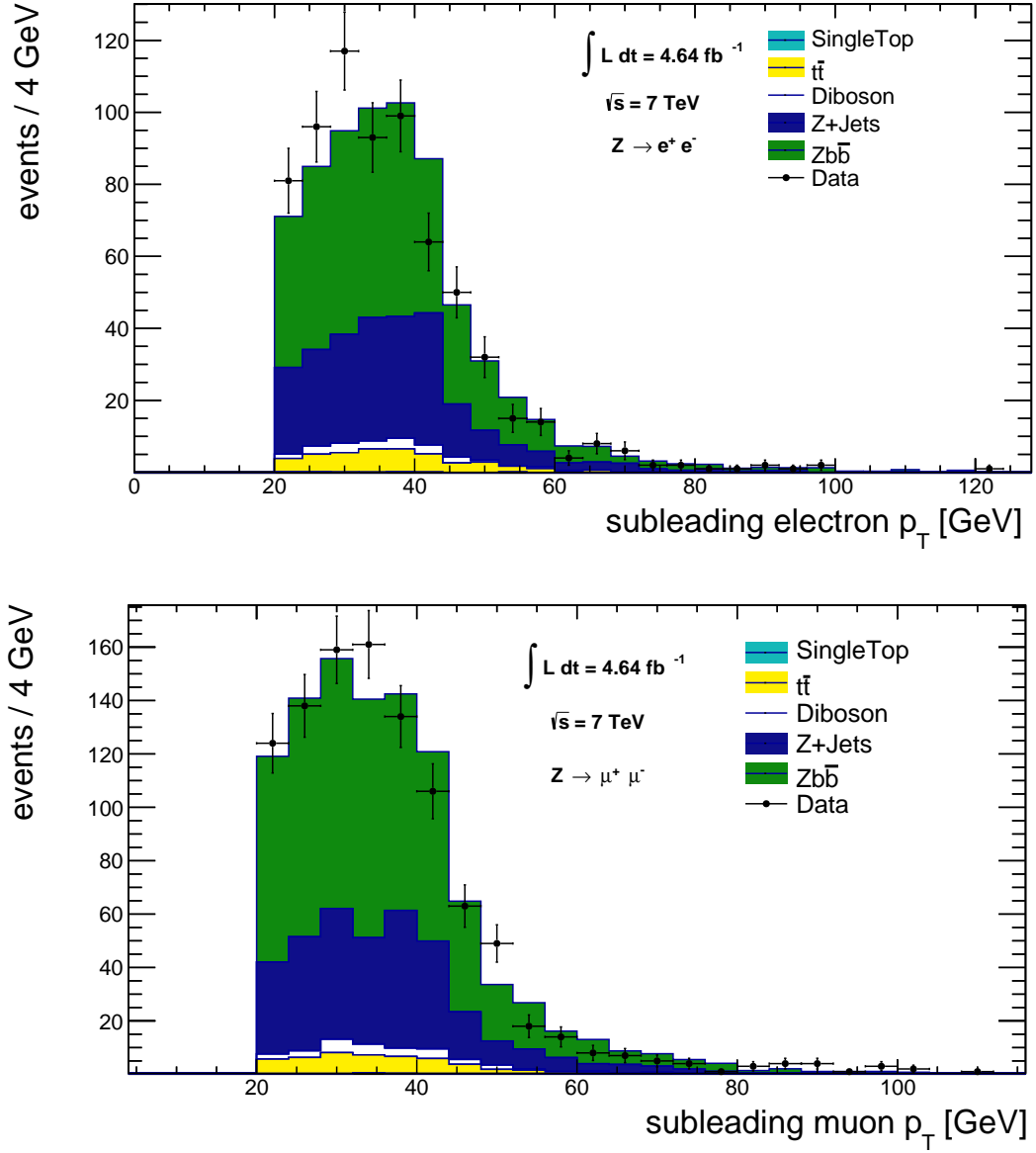


Figure 5.4: Data (points) versus simulation (stacked histograms) distributions of the subleading lepton p_T in the electron channel (top plot) and the muon channel (bottom plot).

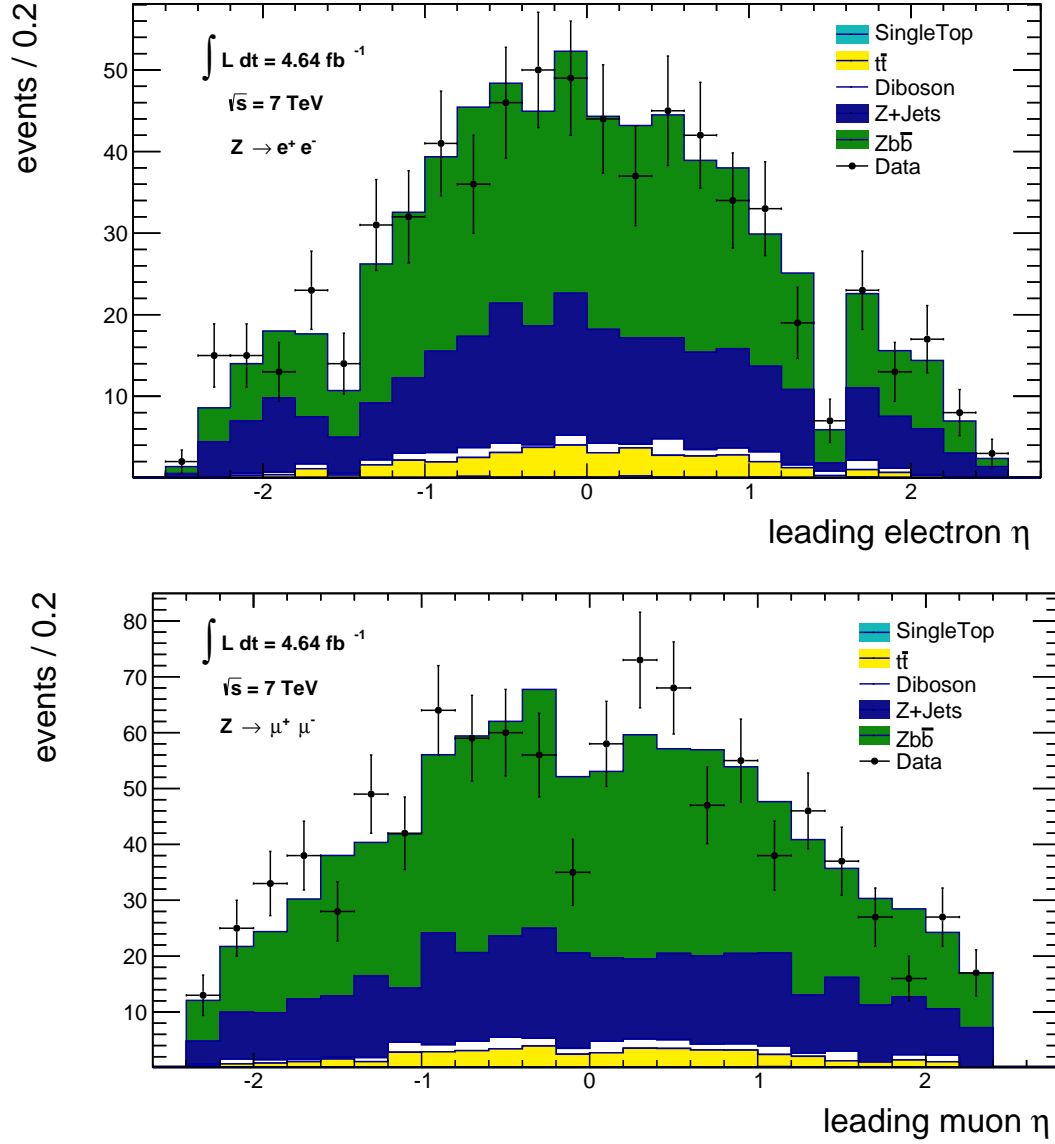


Figure 5.5: Data (points) versus simulation (stacked histograms) distributions of the leading lepton η in the electron channel (top plot) and the muon channel (bottom plot).

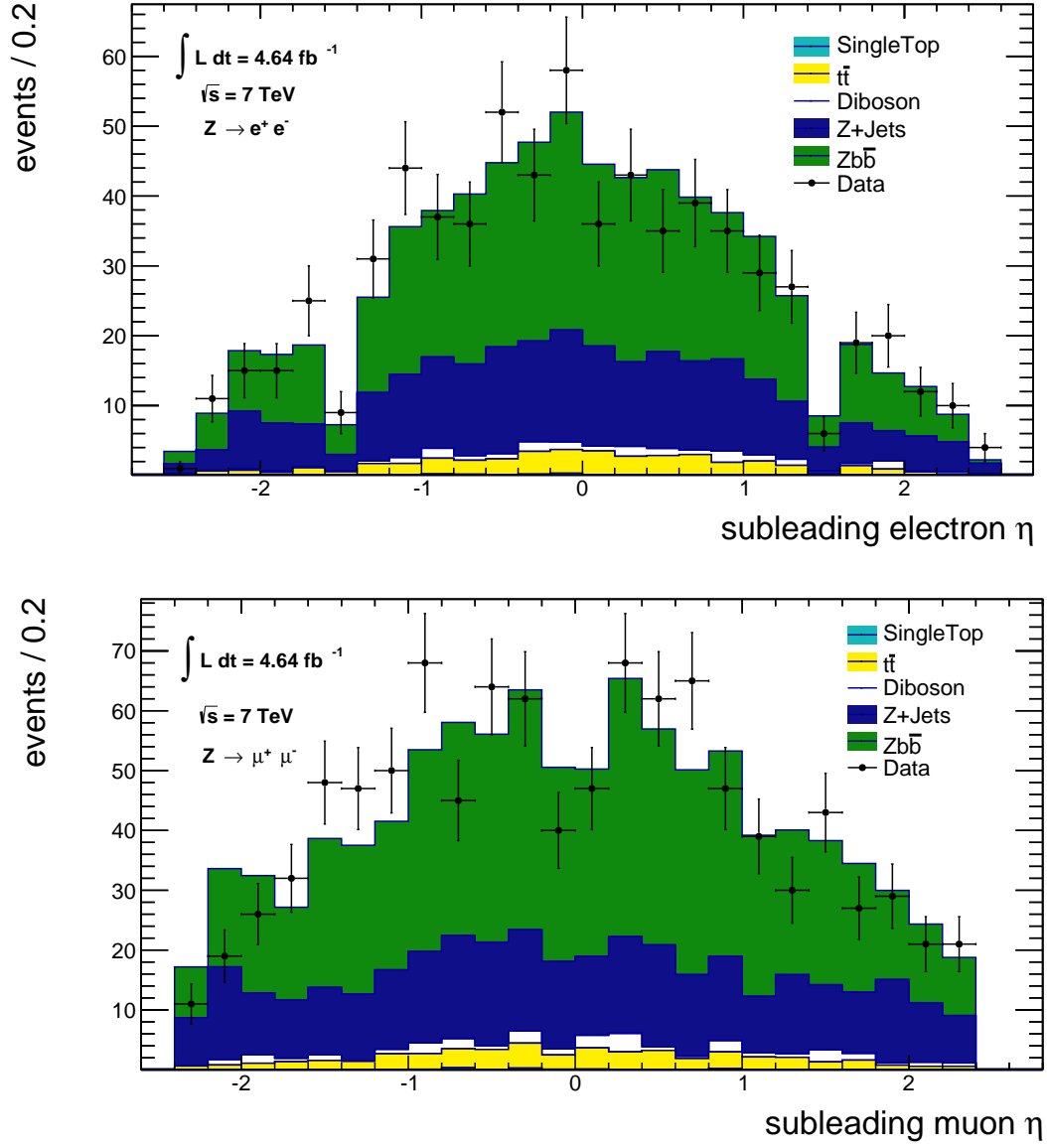


Figure 5.6: Data (points) versus simulation (stacked histograms) distributions of the subleading lepton η in the electron channel (top plot) and the muon channel (bottom plot).

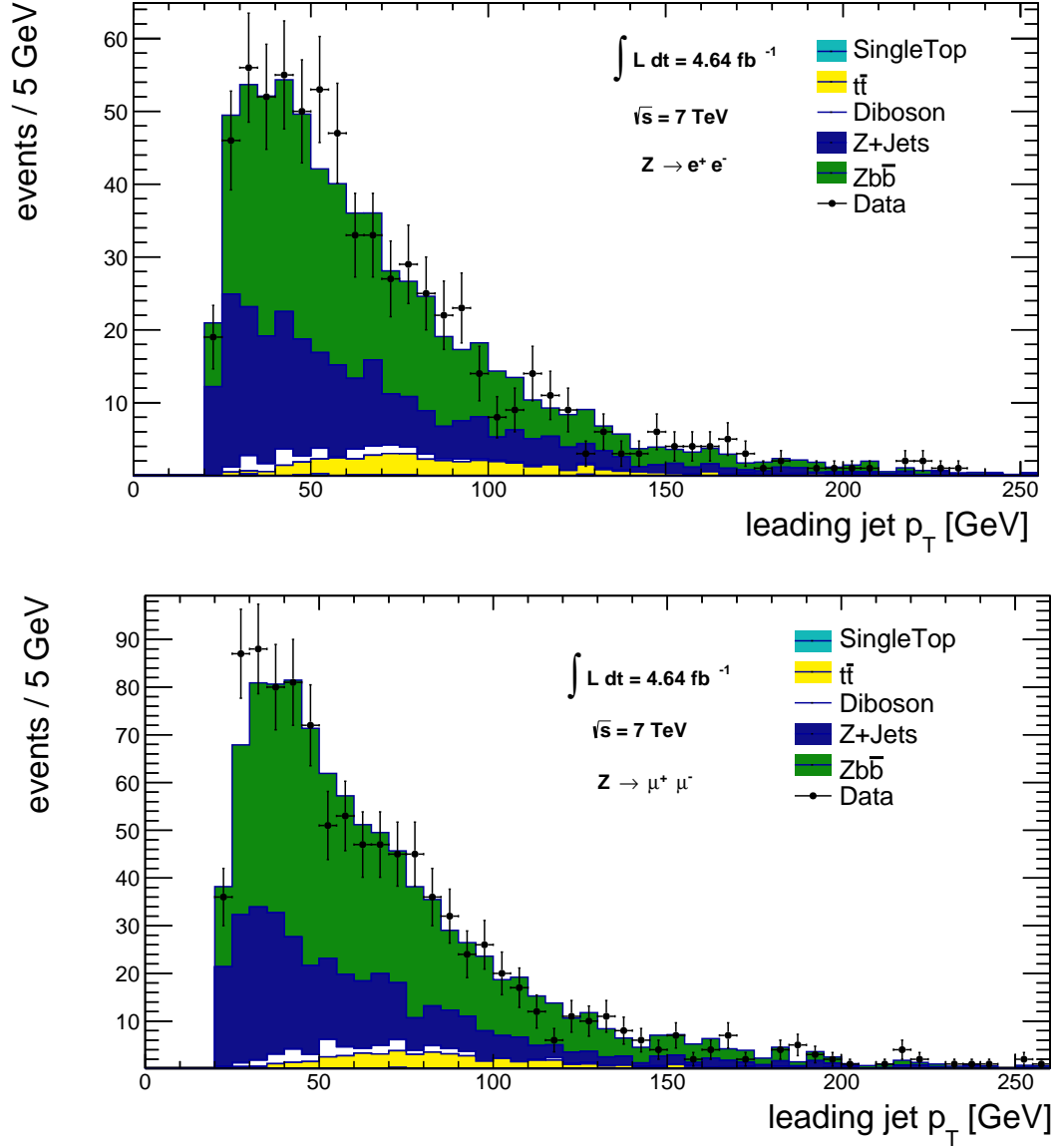


Figure 5.7: Data (points) versus simulation (stacked histograms) distributions of the leading jet p_T in the electron channel (top plot) and the muon channel (bottom plot).

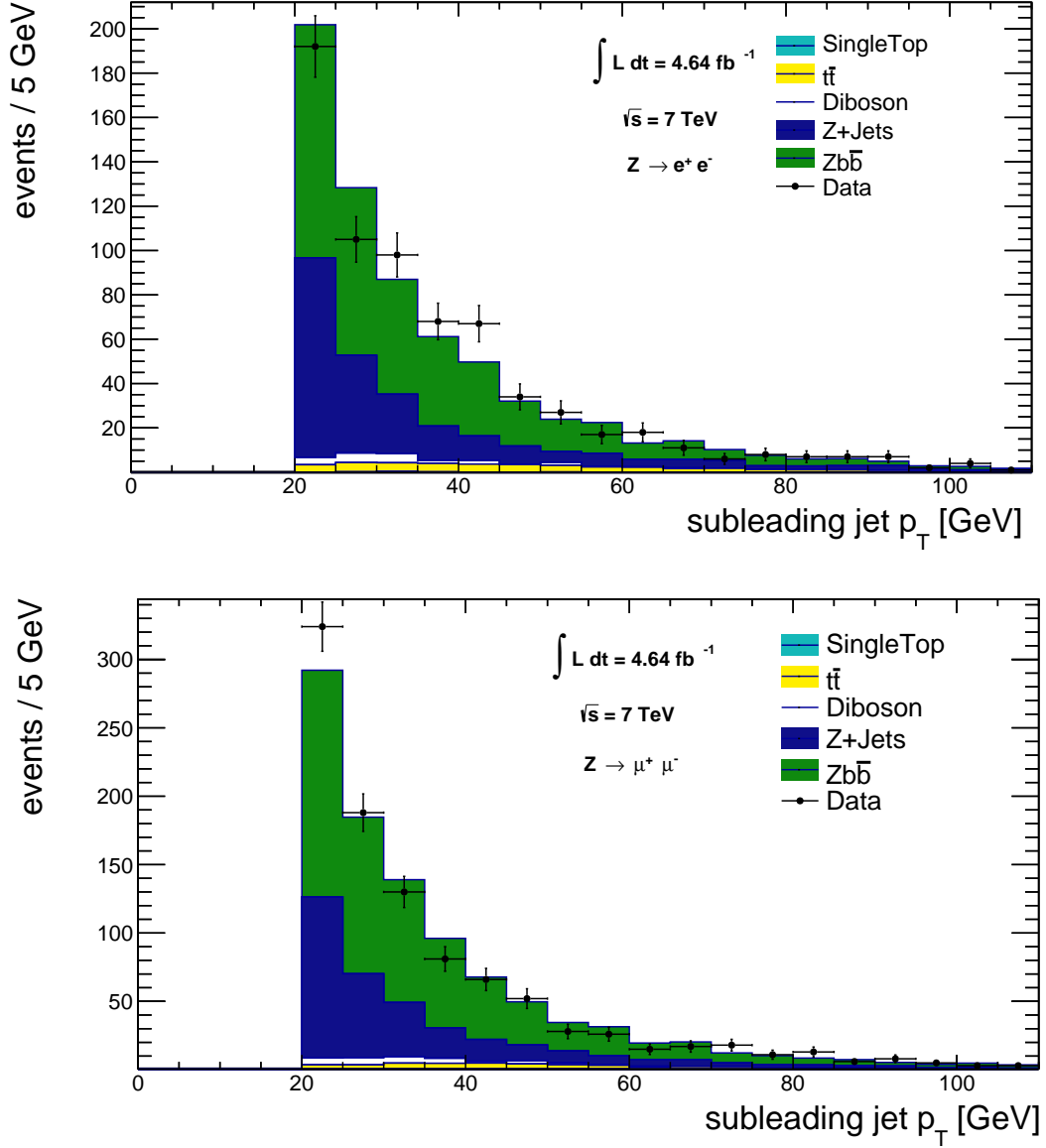


Figure 5.8: Data (points) versus simulation (stacked histograms) distributions of the subleading jet p_T in the electron channel (top plot) and the muon channel (bottom plot).

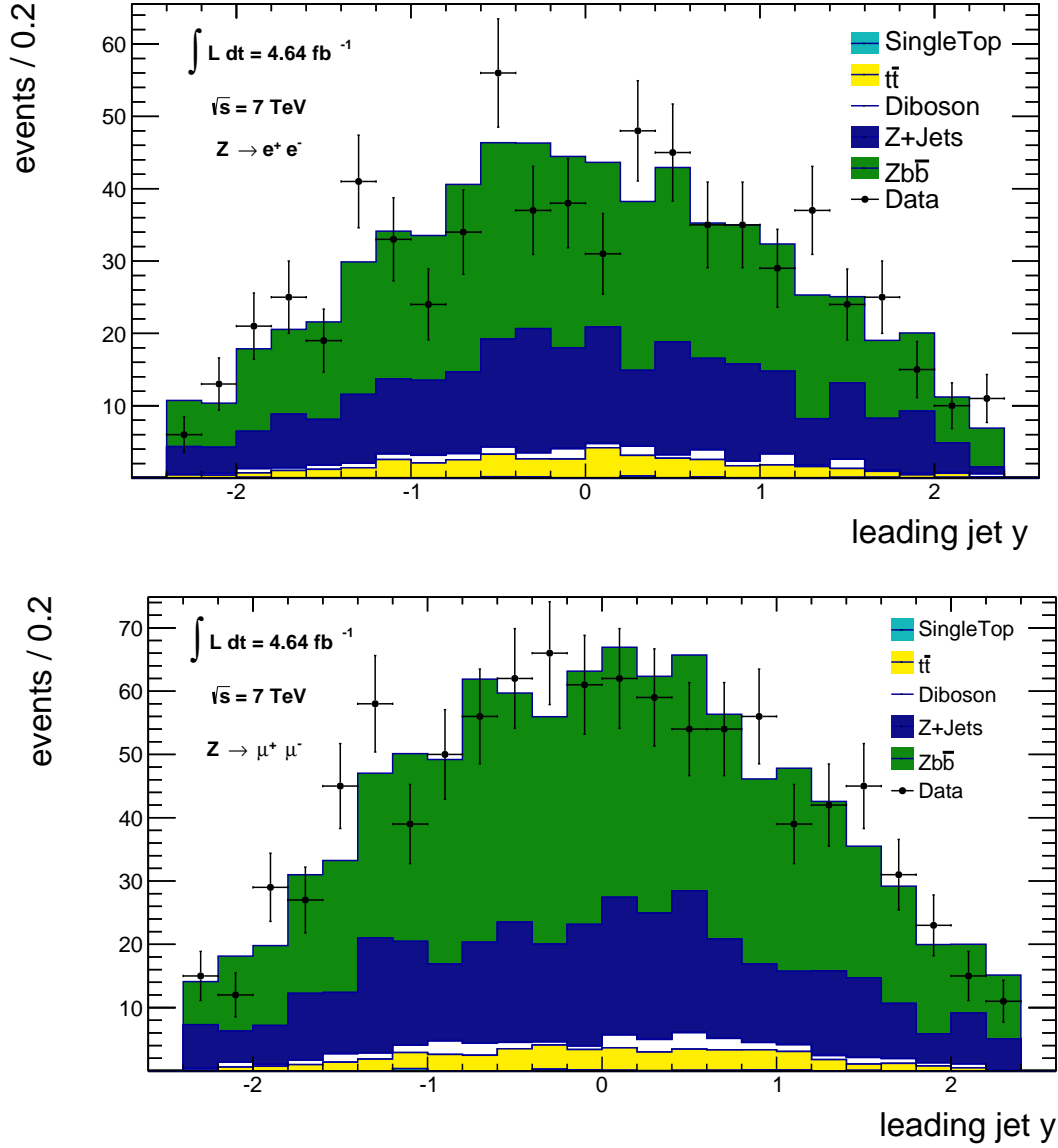


Figure 5.9: Data (points) versus simulation (stacked histograms) distributions of the leading jet rapidity (y) in the electron channel (top plot) and the muon channel (bottom plot).

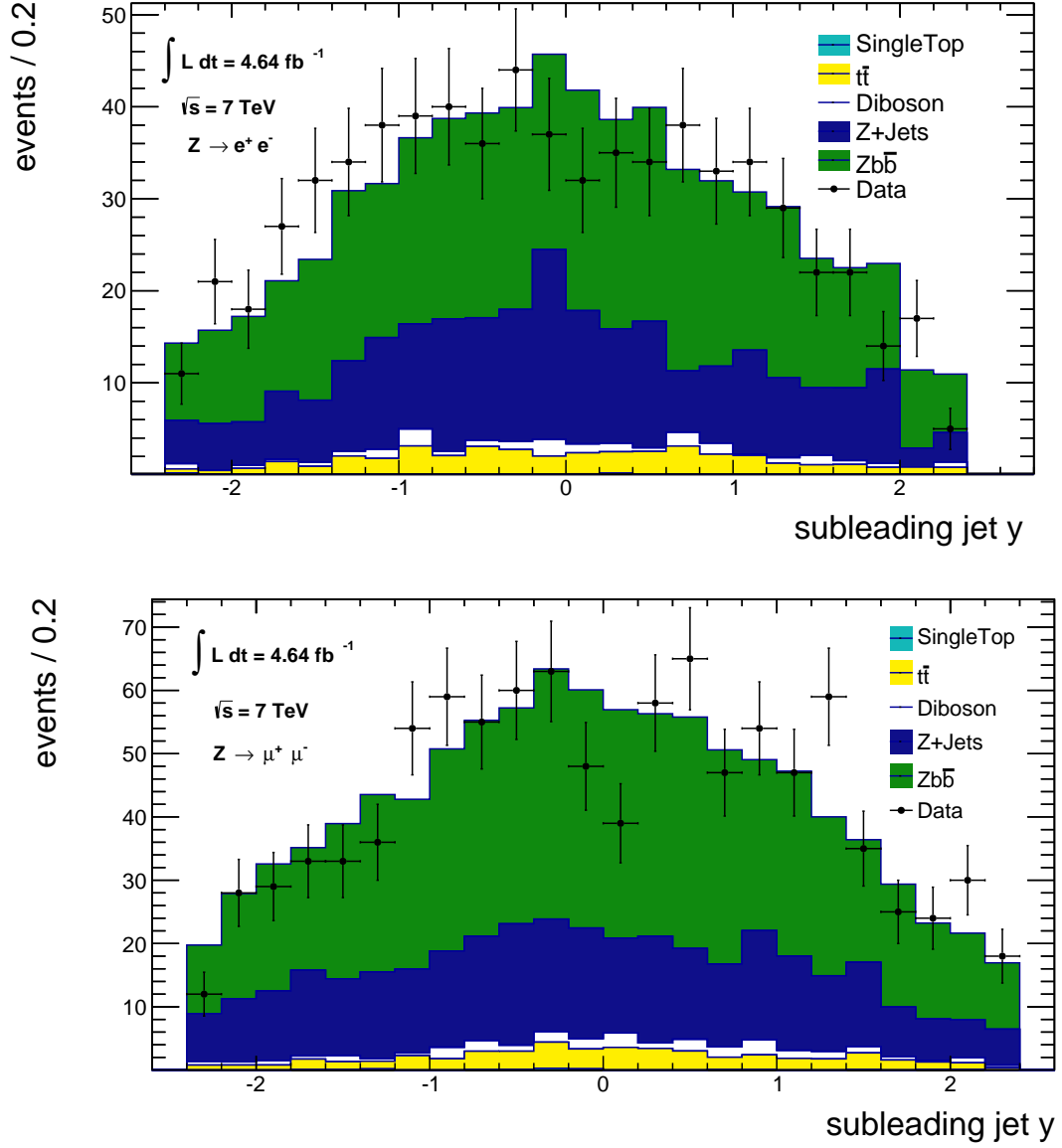


Figure 5.10: Data (points) versus simulation (stacked histograms) distributions of the subleading jet rapidity (y) in the electron channel (top plot) and the muon channel (bottom plot).

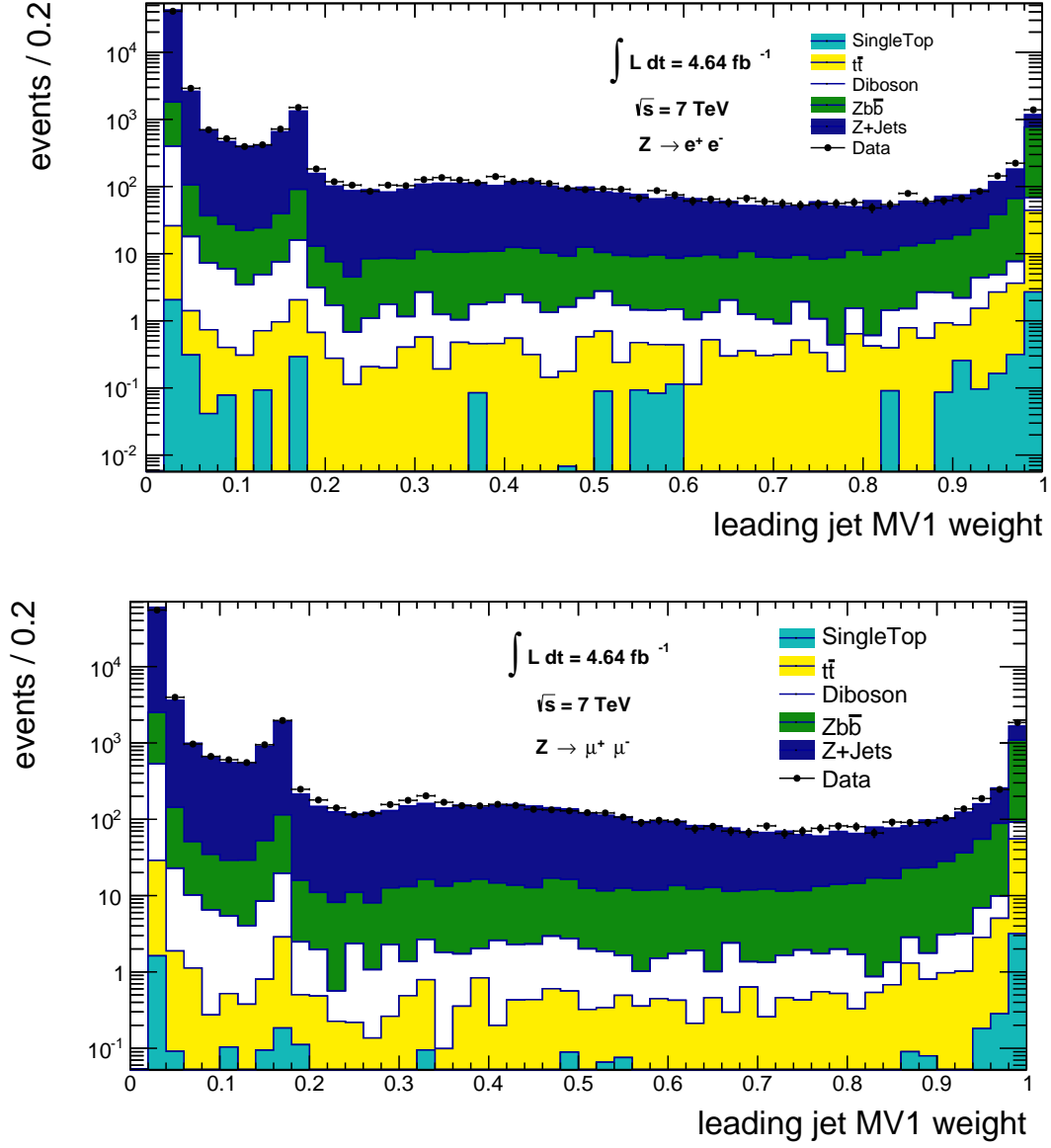


Figure 5.11: Data (points) versus simulation (stacked histograms) distributions of the leading jet MV1 weight before the b -tagging selection in the electron channel (top plot) and the muon channel (bottom plot).

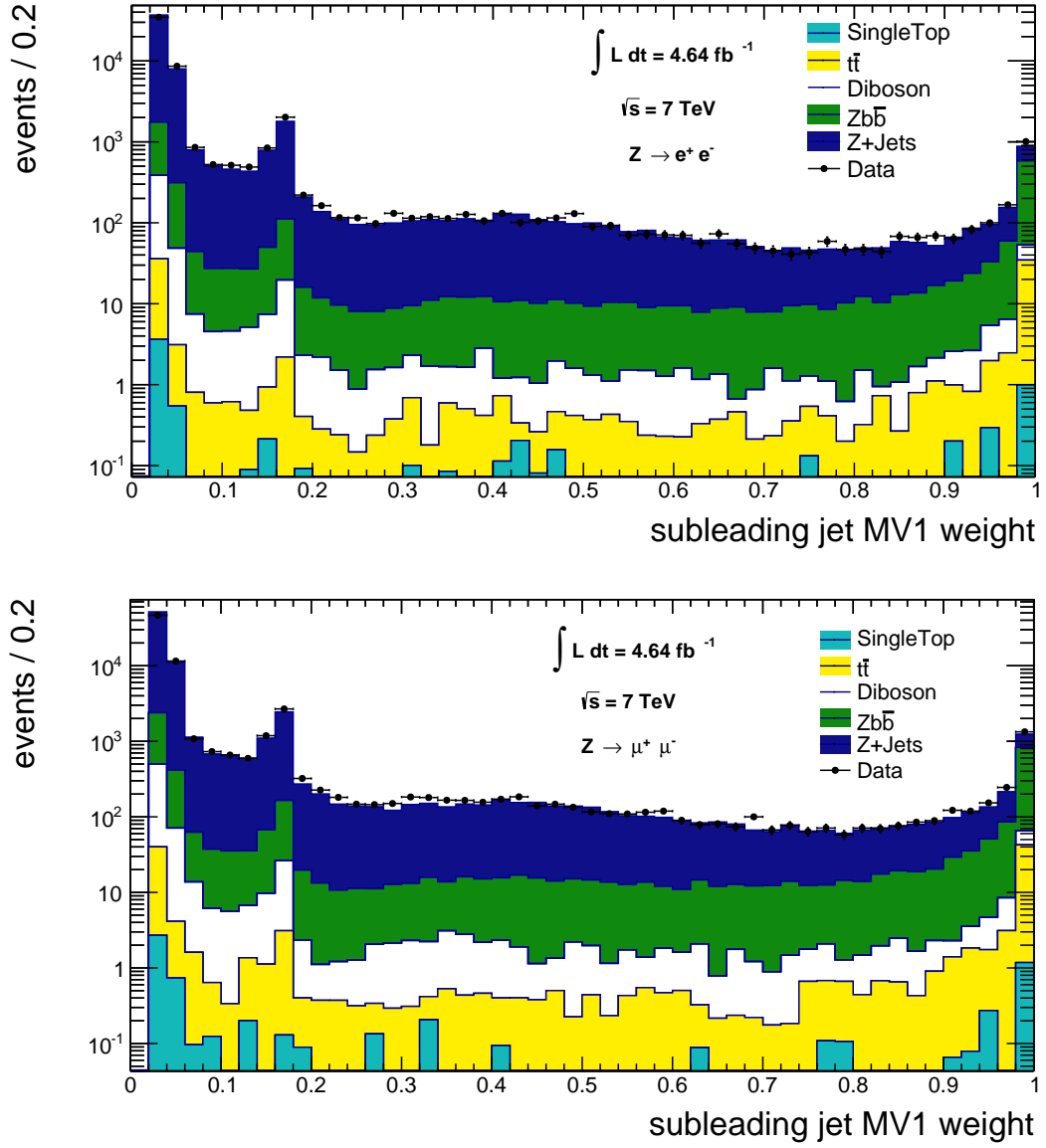


Figure 5.12: Data (points) versus simulation (stacked histograms) distributions of the subleading jet MV1 weight before the b -tagging selection in the electron channel (top plot) and the muon channel (bottom plot).

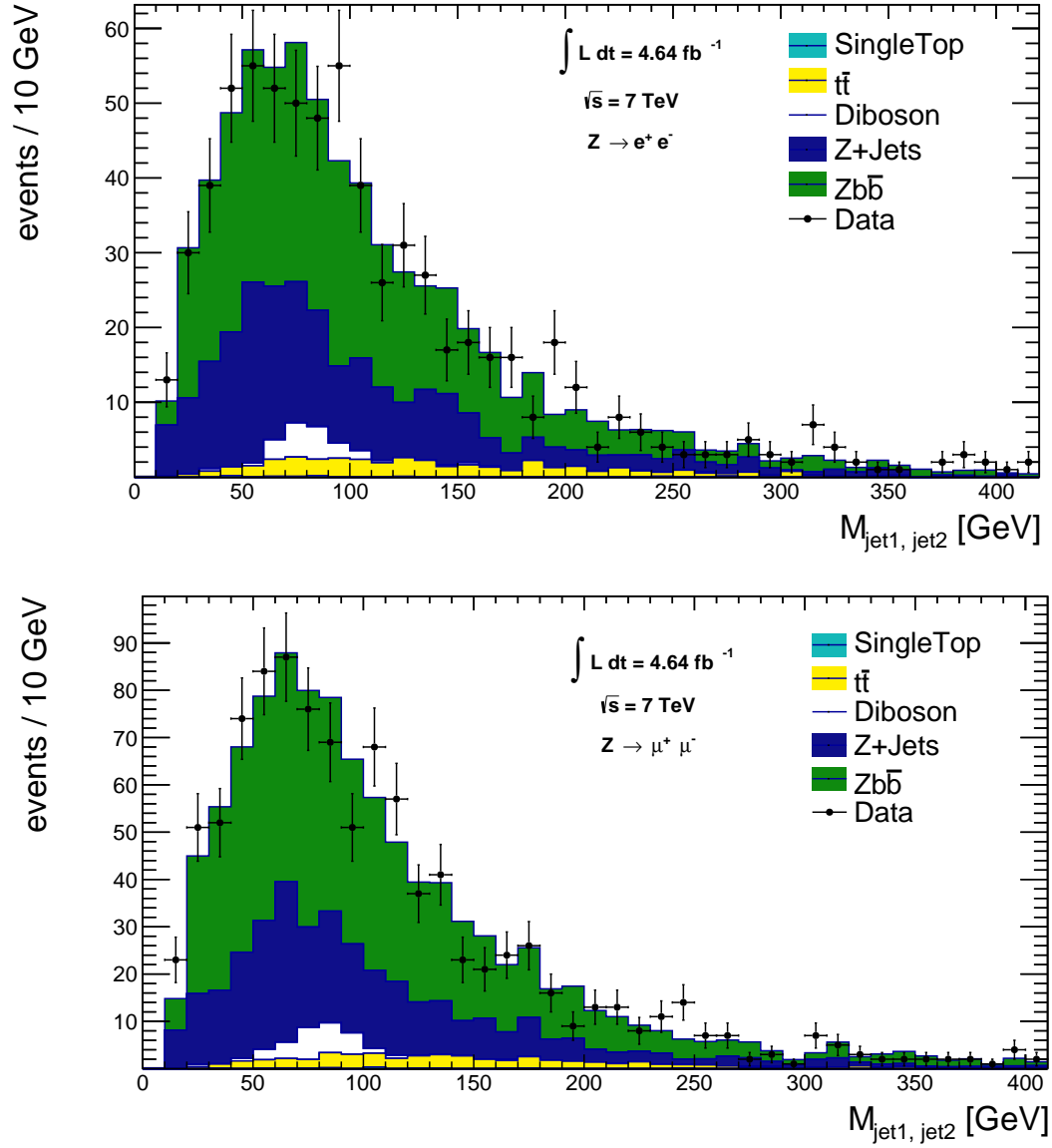


Figure 5.13: Data (points) versus simulation (stacked histograms) distributions of the invariant mass of leading two jets in the electron channel (top plot) and the muon channel (bottom plot).

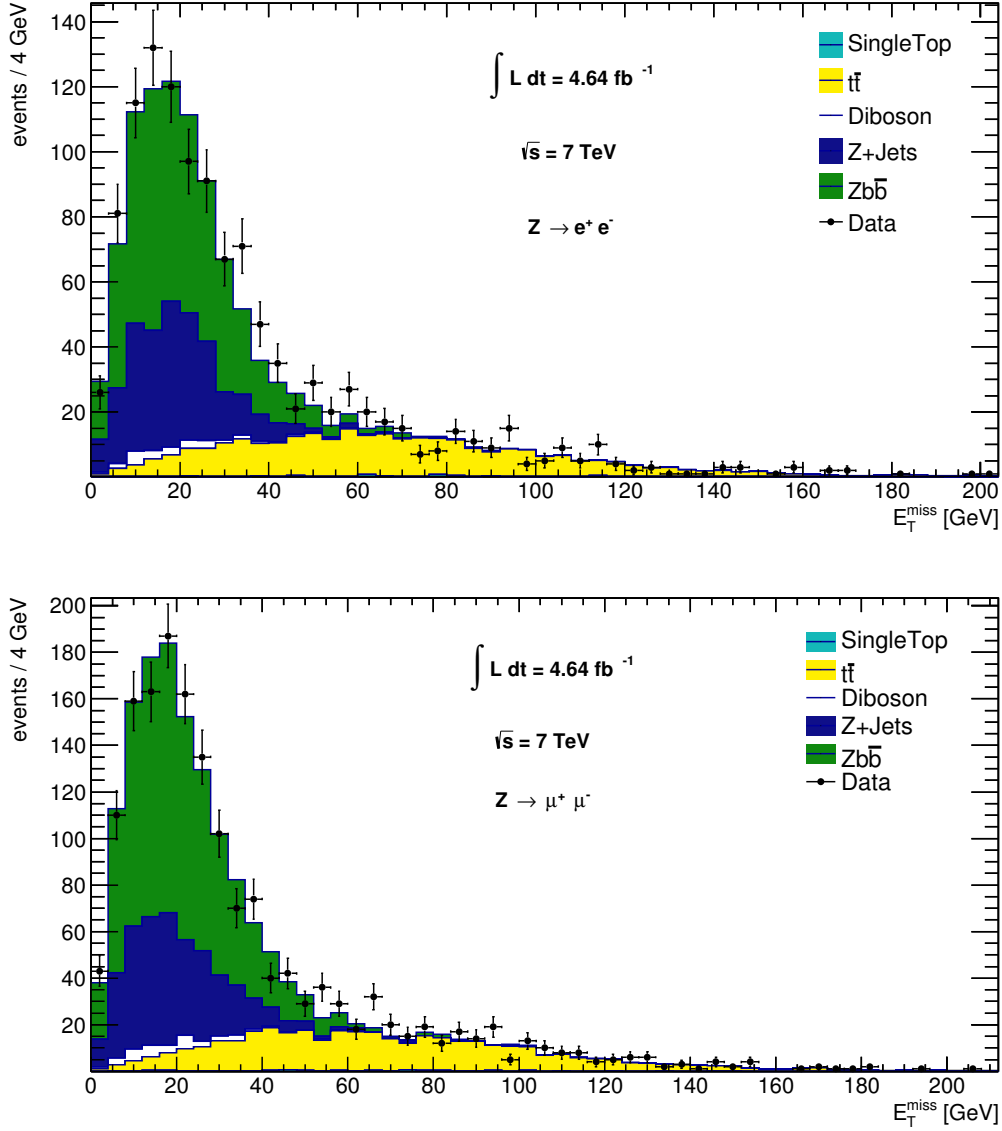


Figure 5.14: Data (points) versus simulation (stacked histograms) distributions of the missing transverse energy before the E_T^{miss} cut in the electron channel (top plot) and the muon channel (bottom plot).

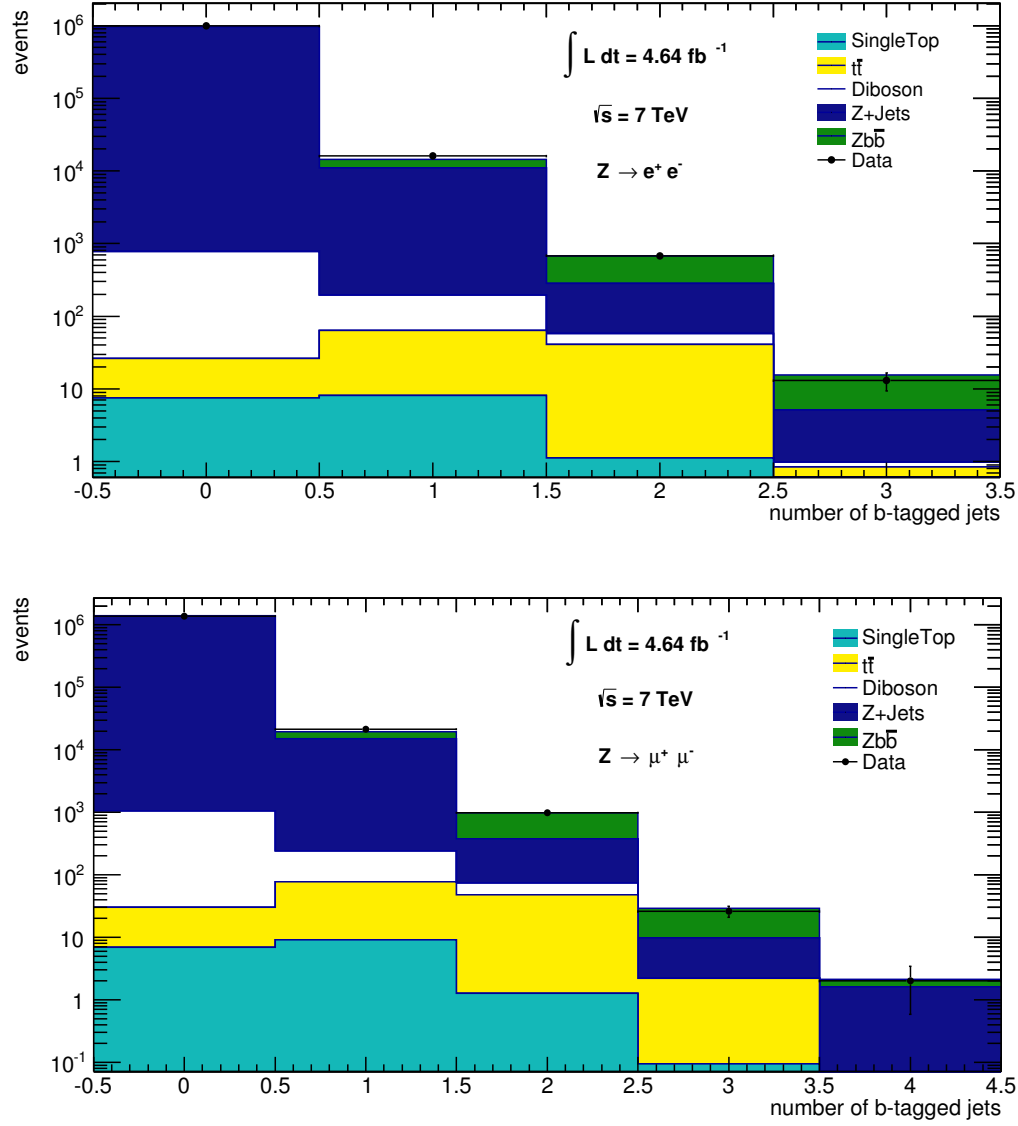


Figure 5.15: Data (points) versus simulation (stacked histograms) distributions of the b -tagged jets multiplicity in the electron channel (top plot) and the muon channel (bottom plot).

Sample ID	Physical Process	MC non-scaled	MC scaled	Data
109300	$Zb\bar{b}+0$ part., $Z \rightarrow e^+e^-$	157.8	163.8	
109301	$Zb\bar{b}+1$ part., $Z \rightarrow e^+e^-$	126.2	130.7	
109302	$Zb\bar{b}+2$ part., $Z \rightarrow e^+e^-$	63.8	66.1	
109303	$Zb\bar{b}+3$ part., $Z \rightarrow e^+e^-$	39.1	40.5	
107650	$Z+0$ part., $Z \rightarrow e^+e^-$	35.0	36.1	
107651	$Z+1$ part., $Z \rightarrow e^+e^-$	30.7	31.6	
107652	$Z+2$ part., $Z \rightarrow e^+e^-$	74.8	76.8	
107653	$Z+3$ part., $Z \rightarrow e^+e^-$	54.6	56.1	
107654	$Z+4$ part., $Z \rightarrow e^+e^-$	20.6	21.1	
107655	$Z+5$ part., $Z \rightarrow e^+e^-$	8.9	9.2	
105200	$t\bar{t}$	40.9	40.9	
117360	<i>Single-top</i> , ev t-channel	0.0	0.0	
108343	<i>Single-top</i> , ev s-channel	0.0	0.0	
108346	Wt inclusive	1.1	1.1	
105930	$ZZ, \ell\ell qq$	15.4	15.4	
105942	$W^+Z, qql\ell$	0.8	0.8	
105972	$W^-Z, qql\ell$	1.1	1.1	
Total Events	—	670.9	691.4	692

Table 5.4: Individual simulation sample contribution to the signal region in the electron channel. The first and second columns give the Sample ID and the underlying physical process respectively. The third (fourth) column labeled “MC non-scaled” (“MC scaled”) corresponds to event yields prior (subsequent) to the template fit scaling. The fifth column gives the data event yield.

Sample ID	Physical Process	MC non-scaled	MC scaled	Data
109305	$Zb\bar{b}+0$ part., $Z \rightarrow \mu^+\mu^-$	203.7	262.8	
109306	$Zb\bar{b}+1$ part., $Z \rightarrow \mu^+\mu^-$	156.8	196.6	
109307	$Zb\bar{b}+2$ part., $Z \rightarrow \mu^+\mu^-$	86.1	106.5	
109308	$Zb\bar{b}+3$ part., $Z \rightarrow \mu^+\mu^-$	48.7	58.9	
107660	$Z+0$ part., $Z \rightarrow \mu^+\mu^-$	48.5	53.1	
107661	$Z+1$ part., $Z \rightarrow \mu^+\mu^-$	37.1	39.1	
107662	$Z+2$ part., $Z \rightarrow \mu^+\mu^-$	98.1	103.2	
107663	$Z+3$ part., $Z \rightarrow \mu^+\mu^-$	67.4	71.0	
107664	$Z+4$ part., $Z \rightarrow \mu^+\mu^-$	26.6	28.1	
107665	$Z+5$ part., $Z \rightarrow \mu^+\mu^-$	14.6	15.3	
105200	$t\bar{t}$	48.4	48.4	
117361	$Single-top, \mu\nu$ t-channel	0.0	0.0	
108344	$Single-top, \mu\nu$ s-channel	0.0	0.0	
108346	Wt inclusive	1.4	1.4	
105930	$ZZ, \ell\ell qq$	22.9	22.9	
105942	$W^+Z, qq\ell\ell$	2.1	2.1	
105972	$W^-Z, qq\ell\ell$	1.4	1.4	
Total Events	—	863.7	1011	1011

Table 5.5: Individual simulation sample contribution to the signal region in the muon channel. The first and second columns give the Sample ID and the underlying physical process respectively. The third (fourth) column labeled “MC non-scaled” (“MC scaled”) corresponds to event yields prior (subsequent) to the template fit scaling. The fifth column gives the data event yield.

6 Multijet Background Estimation

The multijets background is characterised in general by events with a high jet multiplicity and relatively large cross-sections due to the underlying QCD production process. Multijet events with leptons from semi-leptonic decays inside jets as well as fakes can lead to final states similar to that of the signal. However, since the signal event topology requires two isolated leptons in association with two b -tagged jets, the multijets background is expected to play a negligible role in this analysis.

The available multijets MonteCarlo samples are characterized by low event yields in regions prior to the final event selection and virtually no contribution in the final event selection region. Moreover, the simulation fake rates are known to be mismodelled. As a result, the multijets background has been estimated from data using an “ABCD” method. The method, is based on the *a-priori* assumption that the shape of the multijets background distribution as function of the Z boson candidate invariant mass remains constant throughout the event selection.

The ABCD method requires the use of two uncorrelated selection cuts - “P” and “Q”, each of them having a strong discriminant power between the signal and the multijets background, resulting in four distinct regions with varying fractions of signal and multijets events as shown Figure 6.1.

The number of multijet events inside the four fiducial regions are connected via a simple proportionality relation:

$$\frac{N_{Region A}^{multijets}}{N_{Region B}^{multijets}} = \frac{N_{Region C}^{multijets}}{N_{Region D}^{multijets}} \quad (6.1)$$

The background will need to be measured inside the multijets enriched regions – *Region B*, *Region C* and *Region D* while for the signal-like region – *Region A*, the multijets background is estimated with relation (6.1).

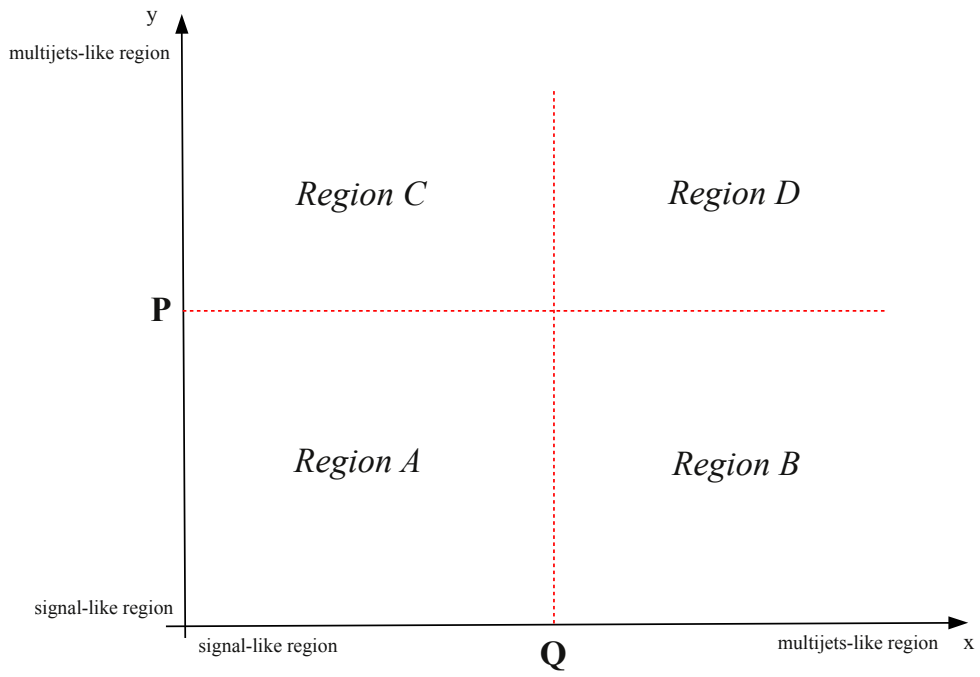


Figure 6.1: Two uncorrelated selections P and Q determine four fiducial regions: Region A (signal-enhanced) and Regions B, C and D (multijet enhanced).

General Method Description

In this section an example for estimating the multijets background in *Region B* as well as some additional details for the estimate inside *Region D* and *Region C* will be given. In this example the “P” selection distinguishes between signal and background leptons and the “Q” selection stands for the entire set of jet selection cuts (see sections 5.1 and 5.2). The four fiducial regions are defined as follows:

- *Region A* – events after jet selection with signal leptons (signal-like region).
- *Region B* – events before jet selection with signal leptons.
- *Region C* – events after jet selection with background leptons.
- *Region D* – events before jet selection with background leptons.

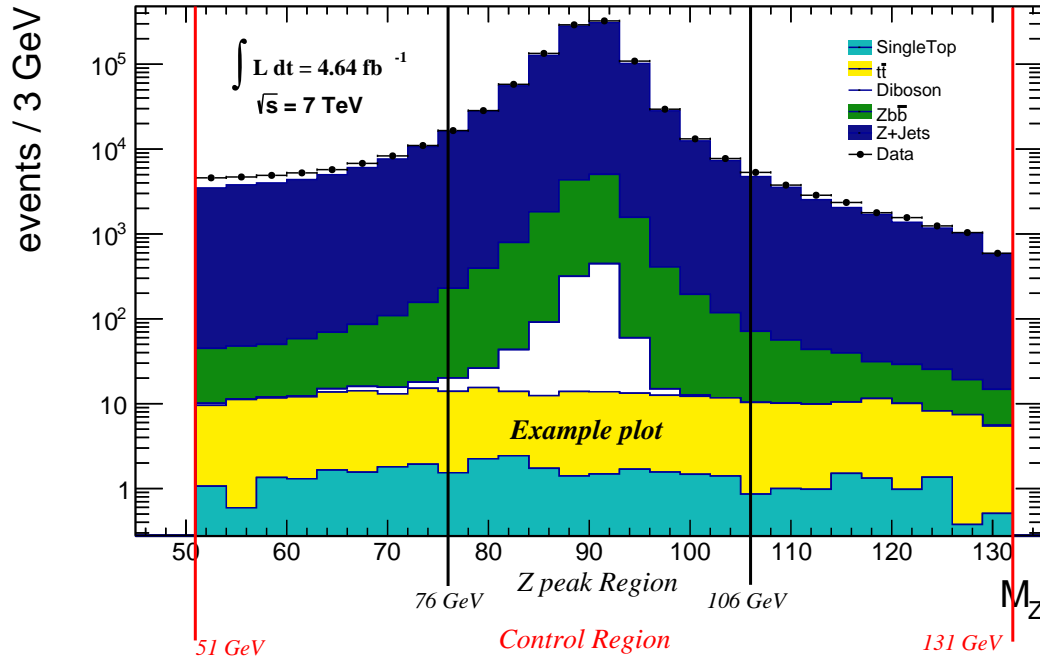


Figure 6.2: Invariant mass distribution of the two signal leptons in data (points) and simulation (histograms) in *Region B*.

Figure 6.2 shows the invariant mass distribution of the two signal leptons corresponding to *Region B* in data and simulation. Two invariant mass windows are chosen as follows:

- a *Z-peak Region* is defined as the invariant mass spectrum corresponding to a window between 76 GeV – 106 GeV.
- a *Control Region* is defined as the invariant mass spectrum corresponding to a window between 51 GeV – 131 GeV.

The estimation of the multijets background in *Region B* is done via a two step approach. First, the simulation distribution is normalized with a scale factor $-f_Z^{Region\ B}$ such that the number of events inside the *Z-peak Region* in simulation matches that in data. This is done in order to account for the missing higher order corrections in simulation. Secondly, the number of scaled simulation events inside the *Control Region* is subtracted from the corresponding region in data. The remaining excess in data is considered to be due to multijets and corresponds to a background estimate outside the *Z-peak Region*¹ but inside the *Control Region*.

In a similar manner, the multijets background is estimated in *Region D*: the simulation is normalized such that it matches the number of events in data inside the *Z-peak Region* and then subtracted from the *Control Region*.

The simulation distribution in *Region C* outside the *Z-peak Region* but inside the *Control Region* is normalized with the scale factor obtained from region *Region D* $-f_Z^{Region\ D}$ and then subtracted from the same distribution in data. The remaining excess of events in data is the estimate of the multijet background in *Region C*.

Finally, by using relation (7.1) the multijets component inside the signal-like region, *Region A* is obtained.

¹Since the number of events in simulation inside the *Z-peak Region* was forced to match that in data, the multijets background in this region will be null.

6.1 Electron Channel

The multijets background in the electron channel has been estimated with the ABCD method by choosing two uncorrelated selections in the following way:

- the product of the charge sign of the two signal electrons will generate two regions, one with $q_1 \cdot q_2 = +1$ (multijet-like region) and one with $q_1 \cdot q_2 = -1$ (signal-like region).
- two additional regions will be generated by applying or not applying the jet selection, a region prior and one following the full jet selection (see Figure 6.3).

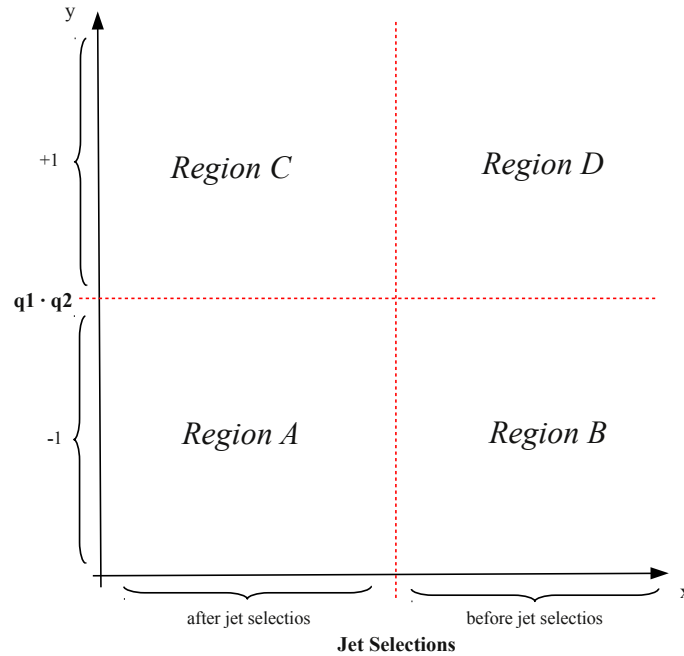


Figure 6.3: Four exclusive regions are generated by the product of the charge sign of the two signal electrons and by applying or not applying the jet selection cuts.

The normalized invariant mass distributions corresponding to the four resulting fiducial regions in data and simulation are shown in Figures 6.4 and 6.5.

For *Region B* and *Region D* the simulation distributions have been normalized such that the number of events inside the *Z-peak Region* equals that in data while the normalization scale factors applied to simulation in *Region C* and *Region A* (after jet selection) have been obtained from the distributions in *Region D* and *Region B* (before jet selection) respectively.

The invariant mass distributions from *Region C* (see Figure , top) and *Region D* (see Figure 6.5) show a peak around the Z-boson even though both fiducial regions correspond to a same-charge sign of the leading and sub-leading electrons. This feature is caused by the limited spatial track resolution inside Inner Detector resulting in a charge misidentification, particularly at high transverse momenta.

Table 6.1 shows the estimated number of multijet events and their statistical uncertainties, corresponding to the four fiducial regions ² outside the *Z-peak Region* but inside the *Control Region*.

The extrapolation inside the *Z-peak Region* in *Region A* is realized by using the multijet distribution after the scaled simulation subtraction from the data in *Region B* as template (see Figure 6.6). The multijets estimation inside the *Z-peak Region*, in *Region B*, was obtained by averaging the excess of events from the left side-band, $QCD_L = 4134$, corresponding to an invariant mass window between 51 GeV – 76 GeV and the right side-band, $QCD_R = 960.54$, corresponding to a window between 106 GeV – 131 GeV from figure 6.6. The average was then rescaled inside the *Z-peak Region*:

$$QCD^{Z\text{-peak}} = \frac{6}{5} \cdot \frac{QCD_L + QCD_R}{2} \quad (6.2)$$

given that the side-bands correspond each to a window of 25 GeV while the *Z-peak Region* corresponds to a 30 GeV window. Since $QCD^B = QCD_L + QCD_R$ the above expression can be written as:

$$QCD^{Z\text{-peak}} = \frac{3}{5} \cdot QCD^B \quad (6.3)$$

The result is that the ratio of multijet events from *Region B* inside the *Z-peak Region* to those outside the *Z-peak Region* but inside the *Control Region* is $3/5$, and is independent of the particular fiducial region studied. Consequently, the multijet background in *Region A* inside the *Z-peak Region* was estimated to 17.9 ± 9.6 (stat.) events (see Appendix A for further details).

² QCD^A corresponds to the multijets estimation from *Region A*, etc.

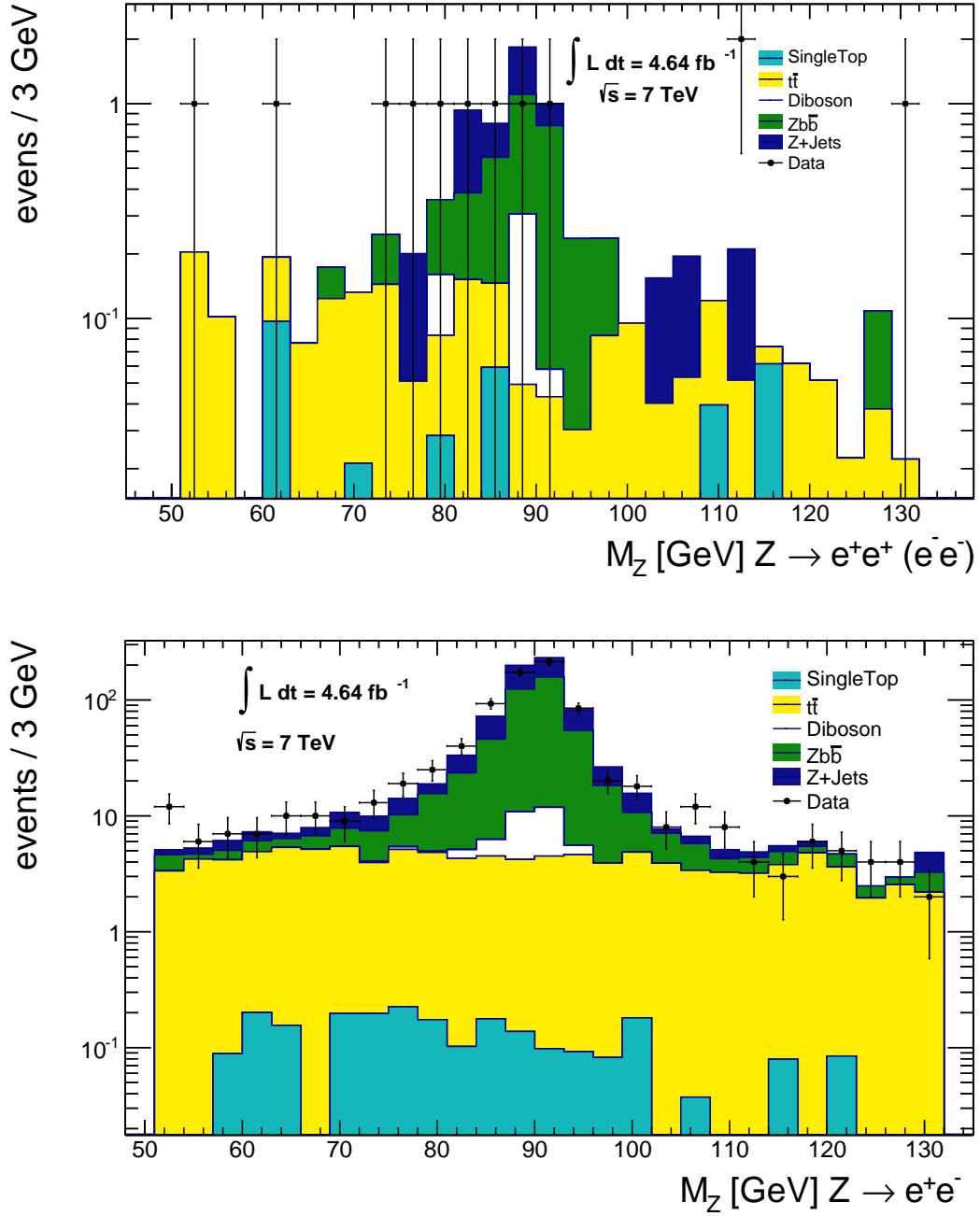


Figure 6.4: Distribution of the invariant mass of the leading and sub-leading electrons for data (points) and the simulation (histograms) after applying the jet selection. The top plot corresponds to the same charge sign (Region C) while the bottom plot corresponds to the opposite charge sign (Region A) of the two electrons. The distribution from Region C (Region A) has been normalized with the scale factor obtained from Region D (Region B).

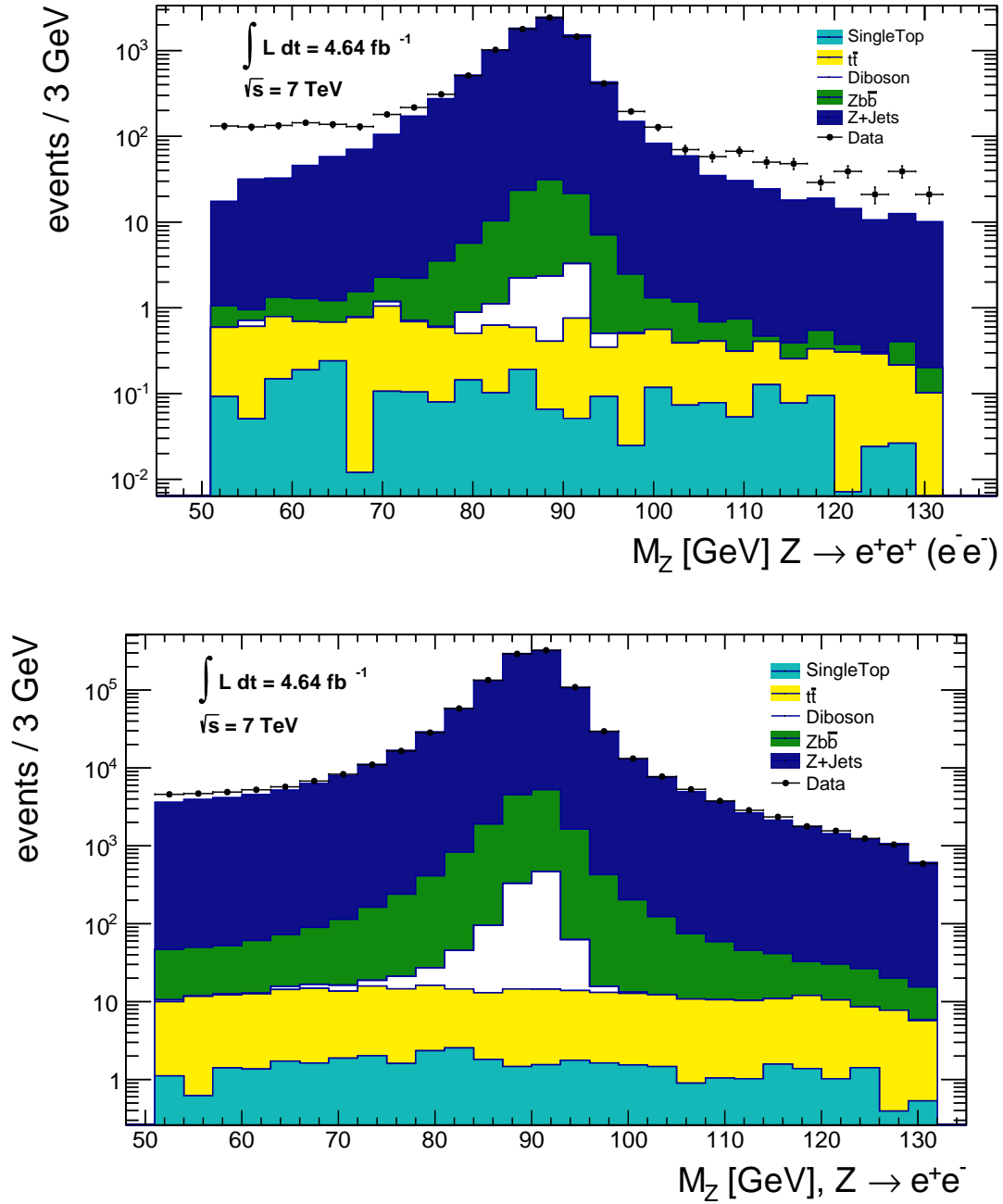


Figure 6.5: Distribution of the invariant mass of the leading and sub-leading electrons for data (points) and the simulation (histograms) before applying the jet selection. The top plot corresponds to the same charge sign (Region D) while the bottom plot corresponds to the opposite charge sign (Region B) of the two electrons. Both simulation distributions have been scaled such that the number of events at the Z-peak equals that in data.

Multijets Region	N_{evts}
QCD^{A}	29.8 ± 16
QCD^{B}	5097 ± 348
QCD^{C}	5 ± 2.6
QCD^{D}	857.4 ± 45.2

Table 6.1: Estimated number of multijet events corresponding the four fiducial regions outside the Z-peak Region but inside the Control Region in the electron channel is shown. The uncertainties are statistical only.

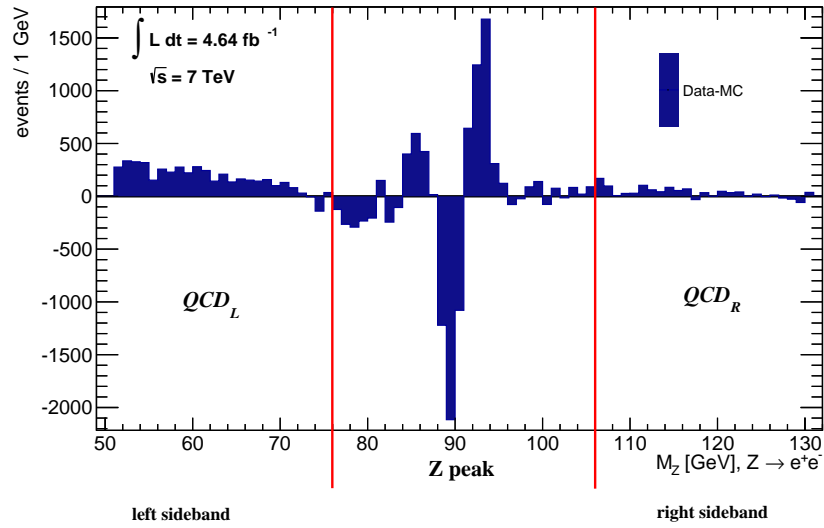


Figure 6.6: The multijets distribution resulted after the scaled simulation subtraction from Region B. The Z-peak Region (76 GeV – 106 GeV), the left side-band (51 GeV – 75 GeV) and right side-band (107 GeV – 131 GeV) are shown.

6.2 Muon Channel

The multijets background in the muon channel has been estimated in a similar way as in the electron channel with the difference that the isolation of the leading and sub-leading muon (see section 5.1) was used instead of the signal leptons charge sign product ³:

- the muon isolation requirement will generate two regions, one dominated by isolated muons (signal dominated) and one dominated by anti-isolated muons (multijet dominated).
- two additional regions are generated by applying or not applying the jet selections, a region prior and one following the jet selection (see figure 6.7).

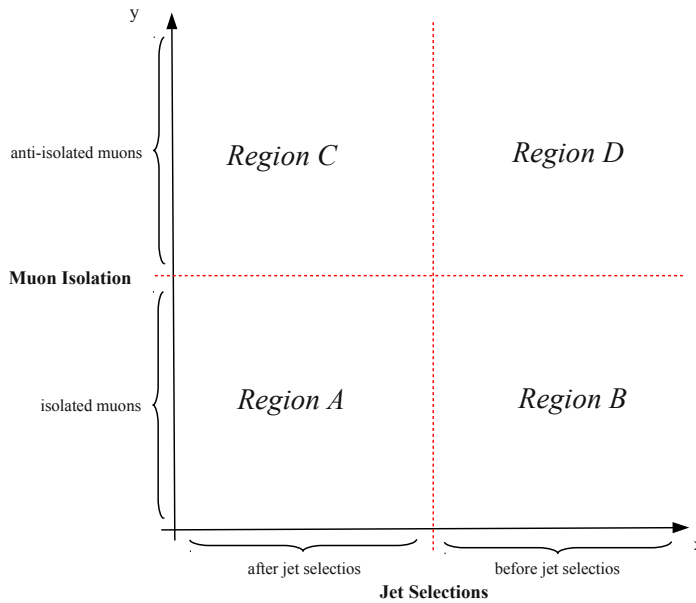


Figure 6.7: Four exclusive regions are generated by requiring isolation or anti-isolation for the two signal muons and by applying or not applying the jet selection cuts.

The corresponding invariant mass distributions of the four fiducial regions are shown in Figures 6.8 and 6.9.

³The signal leptons charge sign product method used in the electron channel has been shown to yield large statistical uncertainties in the muon channel due to the low event yield in the final event selection region. The signal leptons isolation method used in the muon channel could not be applied in the electron channel due to the fact that electrons in the data samples have already been isolated during the di-lepton skimming.

The normalization scale factor applied to simulation in *Region B* (before jet selection) has been obtained by requiring equality between the number of data and simulation events inside the *Z-peak Region*. The same scale factor has been applied to *Region A*.

A large excess of events in data over simulation for the invariant mass distribution in *Region C* and *Region D* can be observed. This is not unexpected since no MonteCarlo samples producing anti-isolated muons significantly were used in this analysis. No normalization factor has been applied to the distributions from *Region C* and *Region D*.

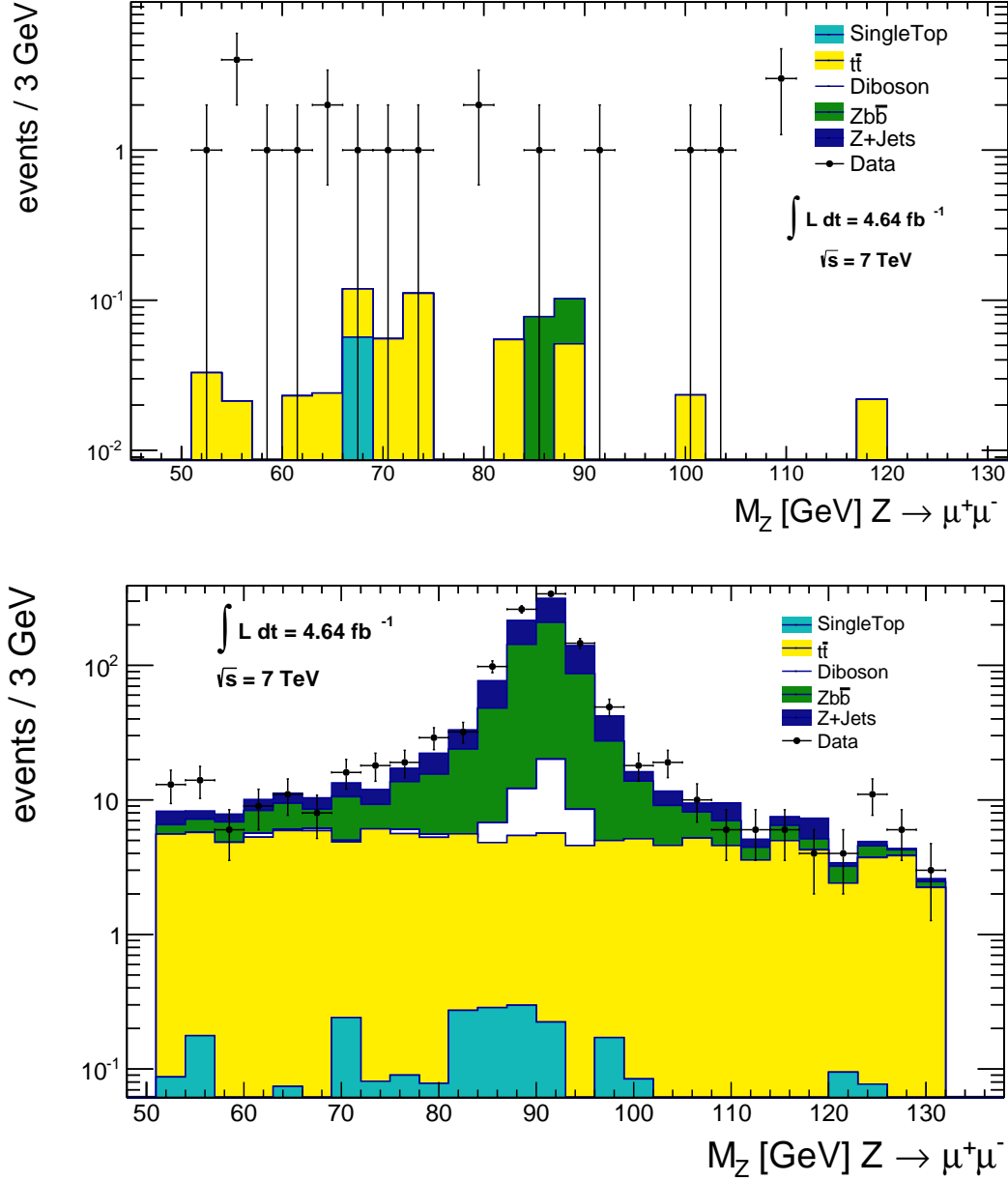


Figure 6.8: Distribution of the invariant mass of the leading and sub-leading muons for data (points) and the simulation (histograms) after applying the jet selection. The top plot corresponds to anti-isolated muons (Region C) while the bottom plot corresponds to isolated muons (Region A). The distribution from Region A has been normalized with the scale factor obtained from Region B.

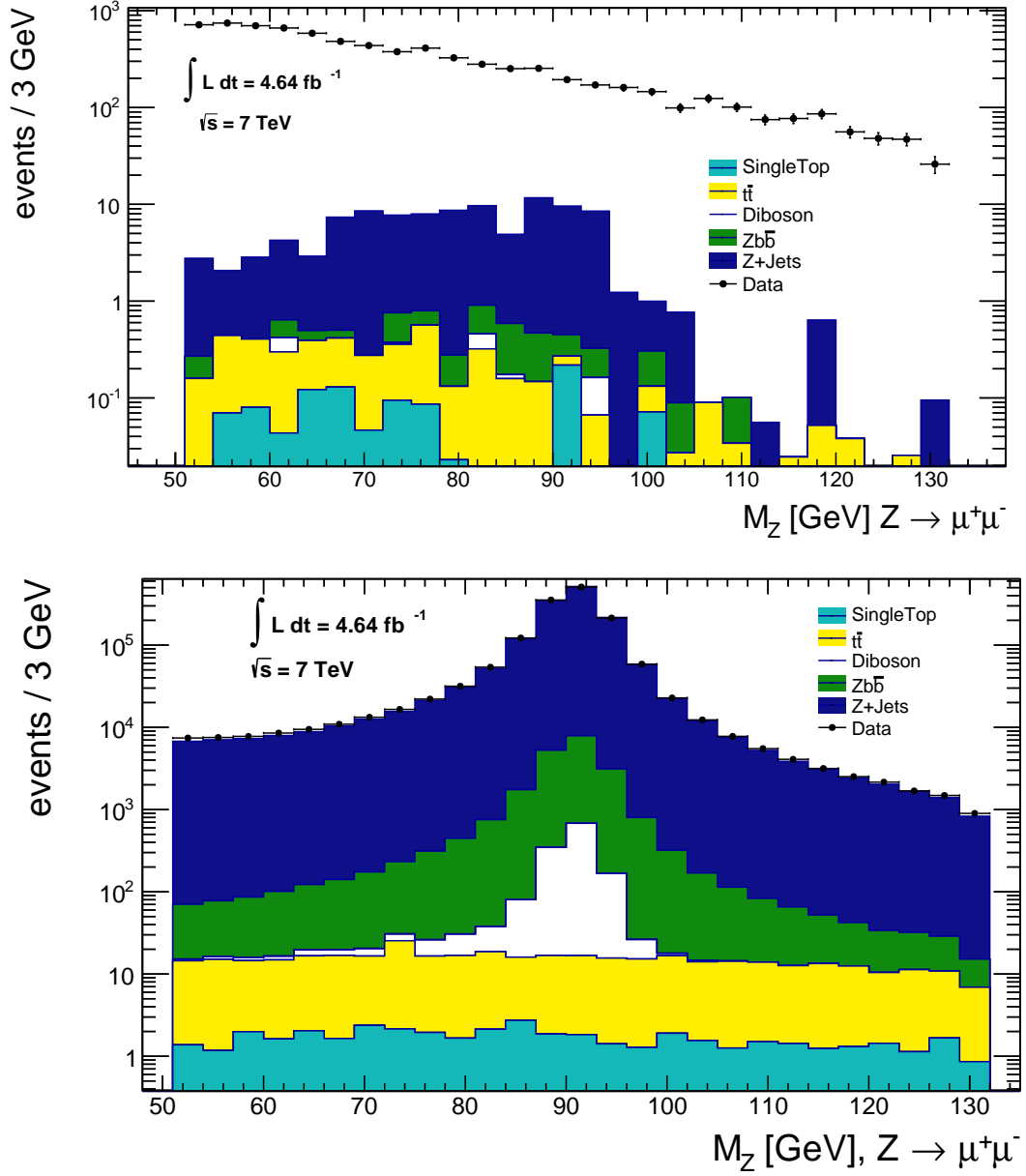


Figure 6.9: Distribution of the invariant mass of the leading and sub-leading muons for data (points) and the simulation (histograms) before applying the jet selection. The top plot corresponds to anti-isolated muons (Region D) while the bottom plot corresponds to isolated muons (Region B). The bottom distribution has been scaled such that the number of events at the Z-peak equals that in data.

Table 6.2 shows the estimated number of the multijets background events and the associated statistical uncertainties corresponding to the four fiducial regions, outside the *Z-peak Region* but inside the *Control Region*.

The multijet distribution after the scaled simulation subtraction from the data in *Region B*, shown in Figure 6.10 was used as template for the background estimation in *Region A*. Just as in the electron channel the left side-band, $\text{QCD}_L = 5303.6$, and right side-band, $\text{QCD}_R = 679.7$, from *Region B* were averaged and the result was used to estimate the multijet contribution inside *Z-peak Region* in *Region B*⁴. The ratio of multijet events inside the *Z-peak Region* to those outside the *Z-peak Region* but inside the *Control Region* in *Region B* was used as a multiplication factor ($3/5$) for the multijets background in *Region A*, QCD^A , in order to estimate the background inside the *Z-peak Region*.

The multijets background estimate is 9.7 ± 2.7 (stat.) events (see Appendix A for further details).

Multijets Region	N_{evts}
QCD^A	16.15 ± 4.5
QCD^B	5977 ± 433
QCD^C	14.6 ± 3.8
QCD^D	5399.16 ± 73.7

Table 6.2: Estimated number of multijet events corresponding the four fiducial regions outside the *Z-peak Region* but inside the *Control Region* in the muon channel is shown. The uncertainties are statistical only.

⁴That is: $\text{QCD}^{\text{Z-peak}} = \frac{3}{5} \cdot (\text{QCD}_L + \text{QCD}_R)$.

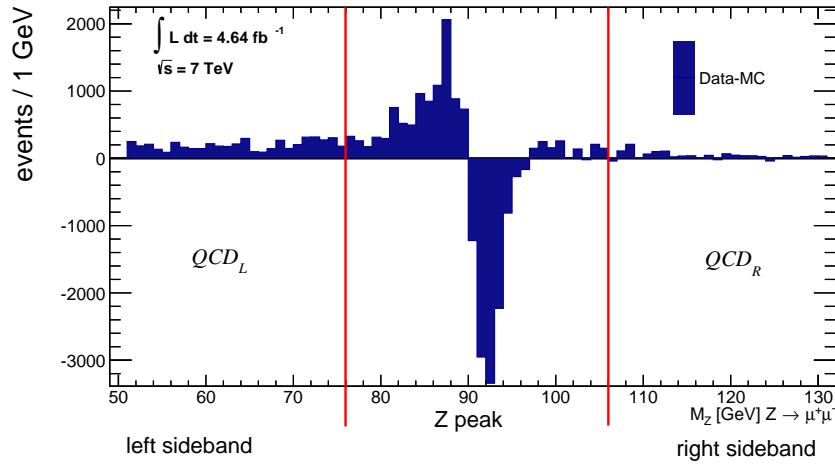


Figure 6.10: The multijets distribution resulted after the scaled simulation subtraction from Region B. The Z-peak Region (76 GeV – 106 GeV), the left side-band (51 GeV – 75 GeV) and right side-band (107 GeV – 131 GeV) are shown.

6.3 Multijets Systematic Uncertainty

To evaluate the systematic uncertainty on the multijet estimation, the background distributions from *Region B*, Figure 6.6 (electron channel) and Figure 6.10 (muon channel) were taken as templates for the multijet evaluation in the signal region (*Region A*). Systematic variations with respect to the central values were obtained as follows. First, each of the two side-bands, the left side-band (for the “up” variation) and right side-band (“down” variation) from *Region B* is rescaled to the *Z-peak Region*:

$$\text{QCD}_{Z\text{-peak}}^{\text{up(down)}} = \frac{6}{5} \cdot \text{QCD}_{L(R)} \quad (6.4)$$

The 6/5 factor comes from rescaling each of the two side-bands (25 GeV window each) to the *Z-peak Region* (30 GeV window). Secondly, the ratio of multijets inside the *Z-peak Region* to that outside the *Z-peak Region* but inside the *Control Region* in *Region B* is assumed to remain constant in the same regions from *Region A*:

$$\text{QCD}_{Z\text{-peak}}^{\text{Region A}} = \frac{\text{QCD}_{Z\text{-peak}}^{\text{up(down)}}}{\text{QCD}_{\text{Region B}}} \cdot \text{QCD}_{\text{Region A}} \quad (6.5)$$

where $\text{QCD}_{\text{Region A(B)}}$ are the multijets background estimates from tables 6.1 and 6.2. In the electron channel the multijet background estimate is:

$$17.9 \pm 9.6 \text{ (stat.)} \pm 11.1 \text{ (syst.) events,} \quad (6.6)$$

while in the muon channel the background has been estimated to:

$$9.7 \pm 2.7 \text{ (stat.)} \pm 7.5 \text{ (syst.) events.} \quad (6.7)$$

Conclusions

A data-driven estimation of the multijet background has been performed for both decay channels, $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$. An excess of the multijet component in the electron channel $17.9 \pm 9.6 \text{ (stat.)} \pm 11.1 \text{ (syst.) events}$ compared to the muon channel $9.7 \pm 2.7 \text{ (stat.)} \pm 7.5 \text{ (syst.) events}$ is observed. Such an asymmetry is expected due to the relatively large (compared to muon fakes) probability of jets to mimic electrons in the EM Calorimeter.

The relatively low multijet event yields compared to those from other processes, confirms the *a-priori* assumption of a negligible background. As a result this background will not be further accounted for in the following.

7 Cross-Section Measurement

In this chapter the extraction of the $Z + b\bar{b}$ cross-section is described. First, the likelihood template-fit method used for the signal fraction extraction from the data sample will be detailed in Section 7.1, followed by a description of the unfolding correction coefficients in Section 7.2. Further, Sections 7.3 and 7.4 will describe the treatment of the systematic uncertainties and the measurement respectively. Finally, Section 7.5 will present a NLO theoretical prediction of the signal.

7.1 Description of the Template Fit

A likelihood template fit approach [64] has been chosen for the extraction of the signal yield in the final event selection region from data, i.e. a reconstructed Z boson in association with at least two b -tagged jets, as described in Section 5.2.

The algebraic sum of a discriminant – “X” of the leading and sub-leading p_T jets, has been chosen as the fit variable. The individual jet’s discriminant – X , is defined as the natural logarithm of the ratio of the jet’s probabilities to have been induced by a b - or c -quark estimated by the JetFitterCOMBNN algorithm:

$$X = \ln(pb/pc)_{\text{leading jet}} + \ln(pb/pc)_{\text{sub-leading jet}} \quad (7.1)$$

The true flavor of each simulated b -tagged jet used in the fit was assessed by matching it to a *bottom* hadron or a *charm* or *light* quark. The matching was done at *particle level* by requiring a weakly decaying b -hadron, with $p_T^{b\text{-hadron}} > 5$ GeV to be within $\Delta R_{\text{jet}, b\text{-hadron}} < 0.3$ to a selected jet. If no b -hadrons were found, a *charm* quark (not originating from a b -hadron decay) was searched for. If neither b -hadrons nor c -quarks were found in the vicinity of a selected jet, the jet was labeled as a *light*-jet. With the definitions given above, six exclusive template distributions become possible depending on the leading and sub-leading jet flavors:

- “ bb ” – both the leading and the sub-leading jets were matched with a b -hadron.

- “ bc ” – the leading (sub-leading) jet was matched with a b -hadron and the sub-leading (leading) jet was matched with a c -quark.
- “ bl ” – the leading (sub-leading) jet was matched with a b -hadron and the sub-leading (leading) jet was matched with a $light$ -quark.
- “ cc ” – both the leading and the sub-leading jets were matched with a c -quark.
- “ cl ” – the leading (sub-leading) jet was matched with a c -quark and the sub-leading (leading) jet was matched with a $light$ -quark.
- “ ll ” – both the leading and the sub-leading jets were matched with a $light$ -quark.

All template distributions are defined at the final event selection stage, i.e. a reconstructed Z boson in association with at least two b -tagged jets.

In order to increase the template stability against statistical fluctuations, the “ bc ”, “ bl ”, “ cc ”, “ cl ” and “ ll ” contributions from the $Z + b\bar{b}$ and $Z + light-partons$ samples (see Tables 4.2 and 4.3) are taken according to their ALPGEN LO cross-sections (scaled to NLO) and were merged into one inclusive template labeled as “ $non-bb$ ” (see Figure 7.1).

Similarly, the “ bb ” contribution is taken exclusively from the $Z + b\bar{b}$ and $Z + light-partons$ samples according to the ALPGEN LO prediction (scaled to NLO) (see Figure 7.1). The normalizations of the bb and $non - bb$ templates are floated as free parameters in the fit.

All flavor contributions from other backgrounds, such as top-antitop, di-boson and single-top, were merged and labeled as “ $other$ ” in the following. They were kept fixed in the fit according to their cross-section predictions (see Table 4.3, Figure 7.1).

The simulation template distributions for “ bb ” - \mathbf{T}^{bb} , “ $non-bb$ ” - \mathbf{T}^{non-bb} and “ $other$ ” - T^{other} are added together and compared with the corresponding distribution observed in data - $T^{observed}$ (see 7.2 top plot). The template fit will change the normalization of the “ bb ” and “ $non - bb$ ” templates until the best agreement with data is reached (see Figure bottom plot):

$$n_{observed} \cdot T^{observed} = \mathbf{n}_{bb} \cdot \mathbf{T}^{bb} + \mathbf{n}_{non-bb} \cdot \mathbf{T}^{non-bb} + n_{other} \cdot T^{other} \quad (7.2)$$

In the above relation $n_{observed}$ represents the number of observed data events while \mathbf{n}_{bb} , \mathbf{n}_{non-bb} are the fitted event yields of the \mathbf{T}^{bb} , \mathbf{T}^{non-bb} templates. The n_{other} event yields of the T^{other} template remains fixed fit.

After performing the template fit to data for the electron, muon and combined channel ¹, the signal $Z + b\bar{b}$ cross-section can be estimated by extracting the signal event yield - \mathbf{n}_{bb} , from the fit:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow e^+ e^-) = \frac{\mathbf{n}_{bb}}{\epsilon_{2b} \cdot C_F \cdot \mathcal{L}}, \quad e \text{ channel} \quad (7.3)$$

¹In the combined channel, both the electron and muon template contributions from simulation are summed and the fit to the data sum of the electron and muon channel templates.

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow \mu^+ \mu^-) = \frac{\mathbf{n}_{bb}}{\epsilon_{2b} \cdot C_F \cdot \mathcal{L}}, \quad \mu \text{ channel} \quad (7.4)$$

and:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+ l^-) = \frac{\mathbf{n}_{bb}}{2 \cdot \epsilon_{2b} \cdot C_F \cdot \mathcal{L}}, \quad \text{combined channel, } l = e, \mu \quad (7.5)$$

In the above relations \mathcal{L} is the integrated luminosity, ϵ_{2b} corresponds to the double-tag efficiency of the MV1 algorithm and C_F is an efficiency correction factor.

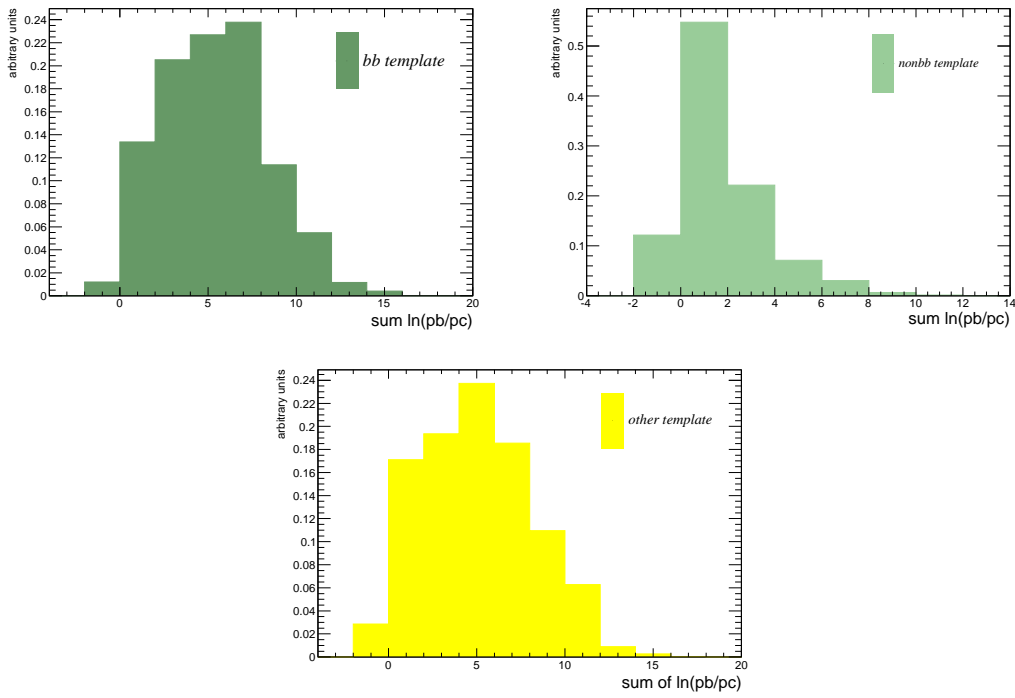


Figure 7.1: Simulation examples for the “bb” (top left plot), “non-bb” (top right plot) and “other” (bottom plot) templates.

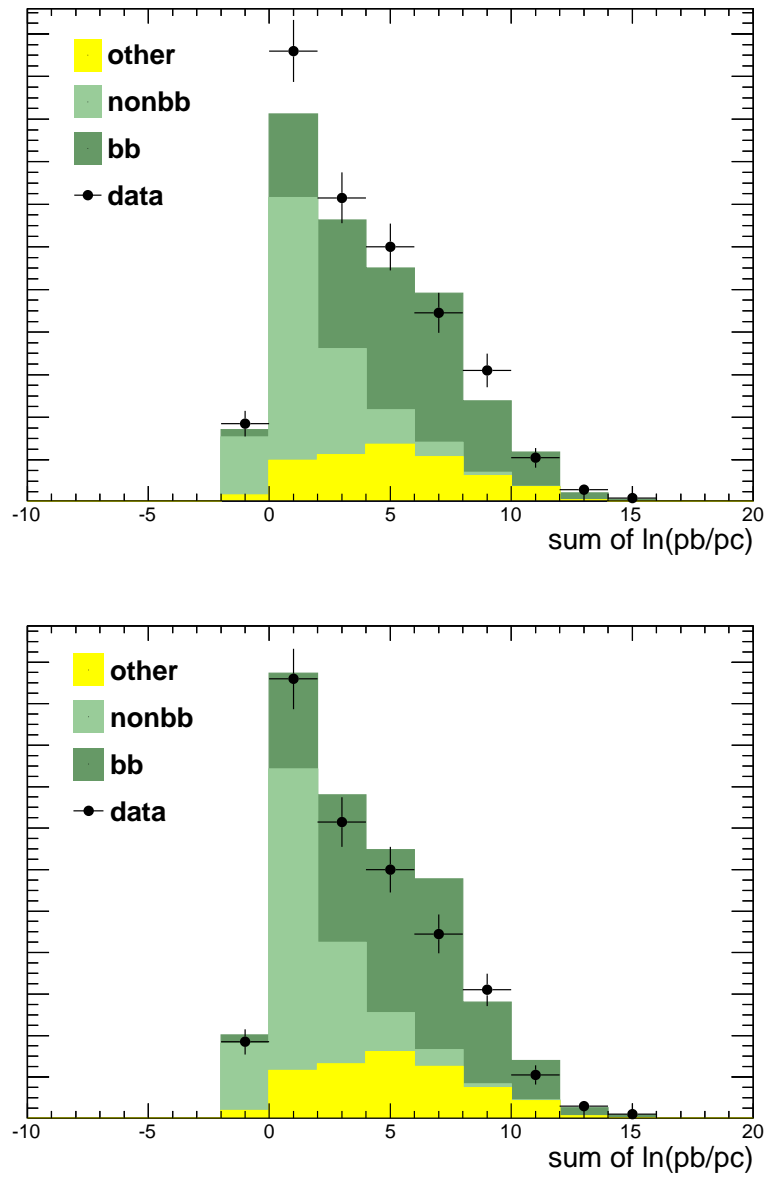


Figure 7.2: Template fit example. Data (dots) versus simulation template distributions before (top plot) and after (bottom) performing the fit.

7.2 Selection Efficiencies Estimation

The \mathbf{n}_{bb} component extracted with the template fit method as described in the previous section, corresponds to an estimation of the signal event yield in the selected data sample with a reconstructed Z boson produced in association with at least two b -tagged jets. Such an estimate is affected by various sources of selection efficiencies, for instance those related to the reconstruction and identification of the physical objects used in the analysis (electrons, muons, jets), the trigger efficiency or the double b -tagging efficiency. A detailed description of the simulation-based methods used in the evaluation of these selection efficiencies is given in the following.

Estimation of the Double b -tagging Efficiency - ε_{2b}

The double b -tagging efficiency, corresponds to the probability of selecting two b -induced jets with the MV1 b -tagging algorithm in the final event selection region. In order to assess this efficiency component, the analysis was run twice on the simulation signal samples using the same set of selection cuts as detailed in Sections 5.1 and 5.2 (including the HFOR decision) with the following differences:

Selection I. The MV1 b -tagging requirement was replaced by forcing each selected jet to be matched with a weakly decaying b -hadron with $p_T^{b-hadron} > 5$ GeV, by requiring $\Delta R_{b-hadron, jet} < 0.3$ to the jet. The fourth row in Table 7.1 (“ $N^{b-matched}$ ”) gives the event yields in the signal region, i.e. a reconstructed Z boson in association with at least two jets matched with b -hadrons.

Selection II. In addition to the selection from **Selection I.**, each jet was also required to have passed the MV1 b -tagging selection. The corresponding event yields in the signal region, are given in the fifth row in Table 7.1, $N_{b-tagged}^{b-matched}$.

The double b -tagging efficiency is given by the ratio of the total event yields:

$$\varepsilon_{2b} = \frac{N_{b-tagged}^{b-matched}}{N^{b-matched}} \quad (7.6)$$

and was estimated to $\varepsilon_{2b}^{comb} = 0.5181 \pm 0.0084$ (stat.) for the combined channel, and $\varepsilon_{2b}^e = 0.5300 \pm 0.0131$ (stat.), and $\varepsilon_{2b}^\mu = 0.5089 \pm 0.0111$ (stat.) for the electron and muon channels respectively.

Estimation of the Efficiency Correction - C_F

For the assessment of the efficiency correction factor – C_F , the analysis was performed at *particle level* following closely the signal event selection described in Sections 5.1 and 5.2 (including the usage of the HFOR tool decision).

To reconstruct the Z boson, the leading and sub-leading *particle level* p_T leptons with opposite charge and of the same flavor were chosen. Furthermore, leptons were selected after the parton shower but before the interaction with the detector material. In order to correct for QED final state radiation effects, *dressed* leptons have been used, i.e. the four-momenta of all photons within $\Delta R_{\text{photon-lepton}} < 0.1$ to a signal lepton were added to the lepton. If both signal leptons were found within $\Delta R_{\text{photon-lepton}} < 0.1$ to a photon, the photon was added to the closest lepton. Finally, only *dressed* leptons with $p_T > 20$ GeV and $|\eta| < 2.5$ were selected. All selected lepton pairs were required to have an invariant mass within $76 \text{ GeV} < M_{ll} < 106 \text{ GeV}$.

Particle level jets were reconstructed passing all particles, including neutrinos, except the signal two *dressed* leptons, to the anti- k_T jet algorithm. All reconstructed jets within $\Delta R_{\text{jet-lepton}} < 0.5$ to a signal lepton were rejected. The selection of b -quark induced jets was done by matching reconstructed jets to weakly decaying b -hadrons using the same procedure as in **Selection I**. Finally, b -jets were required to have $p_T > 20$ GeV and $|y| < 2.4$.

Events containing a reconstructed Z boson in association with at least two truth-jets matched with b -hadrons as described above, were selected.

Table 7.1 shows the signal event yield ($N^{\text{part.level}}$) corresponding to the *particle level* selection as described above.

The efficiency correction factor - C_F is estimated by taking the ratio of the total signal event yield using the selection described in Selection I. to those corresponding to the *particle level* selection:

$$C_F = \frac{N^{b\text{-matched}}}{N^{\text{part.level}}} \quad (7.7)$$

resulting in a value of $C^{comb}_F = 0.3835 \pm 0.0043$ (stat.) for the combined channel, and $C^e_F = 0.3351 \pm 0.0055$ (stat.) and $C^\mu_F = 0.4316 \pm 0.0066$ (stat.) for the electron and muon channels respectively.

Sample ID	Physical Process	$N^{part.level}$	$N^{b-matched}$	$N_{b-tagged}^{b-matched}$
109300	$Zb\bar{b}+0$ part., $Z \rightarrow e^+e^-$	877.42	300.36	154.04
109301	$Zb\bar{b}+1$ part., $Z \rightarrow e^+e^-$	576.71	192.47	101.62
109302	$Zb\bar{b}+2$ part., $Z \rightarrow e^+e^-$	271.397	86.25	46.46
109303	$Zb\bar{b}+3$ part., $Z \rightarrow e^+e^-$	144.42	47.07	26.53
109305	$Zb\bar{b}+0$ part., $Z \rightarrow \mu^+\mu^-$	877.71	386.72	195.32
109306	$Zb\bar{b}+1$ part., $Z \rightarrow \mu^+\mu^-$	568.97	238.02	127.09
109307	$Zb\bar{b}+2$ part., $Z \rightarrow \mu^+\mu^-$	280.66	119.49	64.16
109308	$Zb\bar{b}+3$ part., $Z \rightarrow \mu^+\mu^-$	151.41	63.03	30.49
107650	$Z+0$ part., $Z \rightarrow e^+e^-$	40.68	11.15	9.38
107651	$Z+1$ part., $Z \rightarrow e^+e^-$	11.44	4.88	4.18
107652	$Z+2$ part., $Z \rightarrow e^+e^-$	8.59	2.98	0.62
107653	$Z+3$ part., $Z \rightarrow e^+e^-$	7.38	3.62	1.30
107654	$Z+4$ part., $Z \rightarrow e^+e^-$	4.53	2.02	0.90
107655	$Z+5$ part., $Z \rightarrow e^+e^-$	2.65	0.98	0.42
107660	$Z+0$ part., $Z \rightarrow \mu^+\mu^-$	47.75	17.41	8.86
107661	$Z+1$ part., $Z \rightarrow \mu^+\mu^-$	15.55	11.37	1.13
107662	$Z+2$ part., $Z \rightarrow \mu^+\mu^-$	8.01	4.48	1.46
107663	$Z+3$ part., $Z \rightarrow \mu^+\mu^-$	6.86	3.54	1.50
107664	$Z+4$ part., $Z \rightarrow \mu^+\mu^-$	4.22	2.46	0.94
107665	$Z+5$ part., $Z \rightarrow \mu^+\mu^-$	2.04	0.73	0.16
Total Events	—	3908.43	1499.0	776.58

Table 7.1: The first and second columns represent the MonteCarlo datasets and the corresponding physical process used for the estimation of the efficiency correction factor and double b -tagging efficiency. The third column gives the signal region event yield using the particle level selection. The fourth and fifth columns show the signal region event yields corresponding to selections described in **Selection I** and **Selection II** respectively. All datasets yields have been weighted by cross-section (see Tables 4.2 and 4.3) and luminosity (4.64 fb^{-1}).

7.3 Evaluation of the Systematic Uncertainties

The treatment of systematic uncertainty evaluation will be described in this chapter. The variations with respect to the nominal cross-section result for each systematic uncertainty component has been estimated by applying the procedures detailed in the following, and re-running the event selection. Subsequently, both the template fit “*bb*” signal-fraction and the selection efficiency coefficients were re-estimated. The variation in terms of the signal cross-section for each uncertainty source is given in Table 7.2.

Jet Energy Scale and Resolution

All jet-objects reconstructed with the anti- k_T algorithm are calibrated in a first step at the **EM** scale followed by a second **JES** hadronic scale calibration. An additional jet energy scale correction in bins of p_T and η , accounting for the residual differences between the combined **EM+JES** calibration in simulation and data, has been applied to all jets before the jet selection cuts described in Section 5.1. A systematic uncertainty on the jet calibration correction has been provided by the ATLAS Jet and Missing ET Group. To estimate the impact on the signal measurement, the uncertainty on the residual **JES** correction has been varied by $\pm 1\sigma$ [46].

The electron (muon) channel cross-section variation w.r.t. nominal result from the residual **JES** correction was found to be: $+0.36\%$ (-1.81%) for the $+1\sigma$ (“*up*”) variation, and: -3.69% (-0.02%) for the -1σ (“*down*”) variation. In the combined channel the corresponding up/down variations w.r.t. nominal cross-section result were found to be: -0.95% / -1.51% .

In order to correct for the possible underestimated jet energy resolution (**JER**) in simulation, the four-momentum of each jet, before the kinematic selection but after correcting for the residual **JES** calibration, has been smeared using an analysis package provided by the ATLAS Jet and Missing Et Group.

The variation with respect to the nominal cross-section was symmetrized [61] and was estimated to: $\pm 1.44\%$, $\pm 4.07\%$ and $\pm 3.05\%$ for the electron, muon and combined channel respectively.

B-tagging

A systematic uncertainty on the b -tagging efficiency scale factors (for tagged jets) and the b -tagging inefficiency scale factors (non-tagged jets) corresponding to the MV1 tagger at an 75% operating point (see Section 5.1) is centrally provided by the ATLAS Flavour Tagging Working Group group [52]. The correction scale factors and their corresponding systematic uncertainties present strong p_T , η and jet flavor (b -, c - or *light*-jet) dependencies. In order to estimate the propagation of the correction uncertainties on the signal measurement, the scale factors (SFs) of each of the three exclusive jet flavors have been varied within $\pm 1\sigma$ independently in the following way:

step 1 (step 2). The efficiency scale factors of the b -flavor jets have been varied by $+1\sigma$ (-1σ). Simultaneously the inefficiency scale factors of the b -flavor jets have been varied by -1σ ($+1\sigma$). The up/down variation w.r.t. nominal cross-section result ² in the electron, muon and combined channel was found to be: $-9.20\%/+10.46\%$, $-8.97\%/10.20\%$ and $-9.07\%/+10.29\%$ respectively.

step 3 (step 4). The efficiency scale factors of the c -flavor jets have been varied by $+1\sigma$ (-1σ). Simultaneously the inefficiency scale factors of the c -flavor jets inefficiency SFs have been varied by -1σ ($+1\sigma$). The up/down variation w.r.t. nominal cross-section result in the electron, muon and combined channel was found to be: $+1.91\%/-2.14\%$, $+2.08\%/-2.40\%$ and $+2.04\%/-2.31\%$ respectively.

step 5 (step 6). The efficiency scale factors of the *light*-flavor jets have been varied by $+1\sigma$ (-1σ). Simultaneously the inefficiency scale factors of the *light*-flavor jets have been varied by -1σ ($+1\sigma$). The up/down variation w.r.t. nominal cross-section result in the electron, muon and combined channel was found to be: $+0.07\%/+0.15\%$, $-0.57\%/+0.79\%$ and $-0.32\%/+0.56\%$.

²The up/down variation was taken with respect to the $+1\sigma/-1\sigma$ efficiency SFs ($-1\sigma/+1\sigma$ inefficiency SFs) variation.

HFOR Tool

A systematic uncertainty evaluation of the HFOR approach (described in Section 4.2) can only be estimated for the GS component, common to both Z+light-flavor and Z+heavy-flavor sample-sets. As a result, the evaluation of the systematic uncertainty on the signal measurement introduced by HFOR decision was assessed by selecting both the GF and the GS components from the Z+heavy-flavor sample-sets. Additional events with a heavy quark in the parton shower originating from the Z+light-flavor sample-sets have been rejected.

The variation with respect to the nominal cross-section value has been symmetrized, and was estimated to: $\pm 3.91\%$, $\pm 3.23\%$ and $\pm 3.47\%$ for the electron, muon and combined channel respectively.

Multi-Parton Interactions

Besides the massless bottom (antibottom) quarks produced via gluon splitting processes in the parton shower in ALPGEN Z+light-flavor samples, an additional heavy-flavor quark component with MPI origin can be found in the event. In this particular type of events $b\bar{b}$ quark-pairs can be generated for instance by GS processes from gluons originating from an MPI vertex. Although such events are characterized in general by low b -quark transverse momenta and large pseudorapidities, the possible contribution in the signal region was investigated by scaling up/down by a factor of two/one-half all events which were not flagged by the HFOR tool as heavy-flavored but contained bottom quarks at *particle level*, i.e. events containing exclusively b -quarks with MPI origin.

The up/down variation w.r.t. nominal cross-section result from the MPI contribution in the electron, muon and combined channel was found to be: $-0.24\%/+0.12\%$, $+0.88\%/-0.45\%$ and $+0.38\%/-0.20\%$ respectively.

Electron Corrections

The systematic uncertainty on the electron reconstruction and identification efficiency scale factors, applied to simulation, has been varied independently within $\pm 1\sigma$, using a set of analysis packages provided by the ATLAS ElectronGamma Physics Group [38].

The variation w.r.t. nominal cross-section result from the electron identification efficiency scale factors up/down variation in the electron (combined) channel was found to be: $-2.35\%/+2.43\%$ ($-1.03\%/+1.05\%$).

The variation w.r.t. nominal cross-section result from the electron reconstruction efficiency scale factors up/down variation in the electron (combined) channel was found to be: $-1.90\%/+1.95\%$ ($-0.84\%/+0.85\%$).

In order to estimate the systematic uncertainty related to the electron energy scale corrections in data as well as the four-momentum energy resolution correction (smearing) in simulation, the systematic uncertainty of each correction has been varied within $\pm 1\sigma$ [40].

The variation w.r.t. nominal cross-section result from the electron smearing scale factors up/down variation in the electron (combined) channel was estimated to: $+0.13\%/+0.38\%$ ($+0.06\%/+0.16\%$).

The variation w.r.t. nominal cross-section value from the electron energy scale corrections up/down variation in the electron (combined) channel was found to be: $+0.39\%/-0.44\%$ ($+0.16\%/-0.17\%$).

Muon Corrections

The systematic uncertainty on the muon identification efficiency correction, that on the muon four-momentum smearing and energy scale corrections applied in simulation are provided by the ATLAS Muon Combined Performance Group [43]. Each set of the smearing correction components, i.e. those related to the Muon Spectrometer (MS) and Inner Detector (ID), was varied independently by $\pm 1\sigma$ with respect to its nominal value and the variation with respect to the central cross-section measurement has been evaluated.

The variation w.r.t. nominal cross-section result from the muon identification efficiency up/down variation in the muon (combined) channel was found to be: $-0.76\%/+0.78\%$ ($-0.43\%/+0.43\%$).

The variation in terms of measured cross-section w.r.t. nominal measurement from the MS smearing up/down corrections in the muon (combined) channel was estimated to:

$-0.86\%/-0.09\%$ ($-0.06\%/-0.04\%$).

The variation w.r.t. nominal cross-section measurement from the ID smearing scale factors up/down variation in the muon (combined) channel was found to be: $-0.04\%/+0.02\%$ ($0.02\%/0.01$).

The variation w.r.t. nominal result from the muon energy scale correction up/down variation in the muon (combined) channel was estimated to: $\pm 0.20\%$ ($\pm 0.11\%$).

Missing Transverse Energy

The missing transverse energy systematic uncertainty has been estimated by replacing the reconstructed E_T^{miss} component (accounting for the lepton and jet smearing, jet energy scale correction, etc.) with the uncorrected value, i.e. the original E_T^{miss} component stored in the event.

The variation with respect to the central cross-section value has been symmetrized and was found to be $\pm 2.23\%$ in the electron channel, $\pm 0.29\%$ in the muon channel and $\pm 1.06\%$ in the combined channel.

Lepton Trigger

The lepton trigger scale factors have been varied within $\pm 1\sigma$ independently in each channel and the variation with respect to the nominal cross-section measurement was taken as an estimate of the trigger systematic uncertainty.

The variation in terms of measured cross-section w.r.t. nominal result from the lepton trigger scale factors up/down variation was estimated to be: $-0.06\%/-0.06\%$ in the electron channel, $-0.54\%/+0.57\%$ in the muon channel and $-0.35\%/+0.35\%$ in the combined channel.

Luminosity

The systematic uncertainty on the integrated luminosity of full 2011 data sample used in this analysis was estimated to 1.8% [60].

Syst. component	Electron channel	Muon channel	Combined channel
E_T^{miss}	$\pm 2.2\%$	$\pm 0.29\%$	$\pm 1.0\%$
JER	$\pm 1.4\%$	$\pm 4.0\%$	$\pm 3.1\%$
JES \uparrow	$+0.36\%$	-1.8%	-1.0%
JES \downarrow	-3.7%	-0.02%	-1.5%
B – tagging \uparrow	-9.2%	-9.0%	-9.0%
B – tagging \downarrow	$+10.5\%$	$+10.2\%$	$+10.3\%$
C – tagging \uparrow	$+1.9\%$	$+2.1\%$	$+2.0\%$
C – tagging \downarrow	-2.1%	-2.4%	-2.3%
Light – tagging \uparrow	$+0.07\%$	-0.6%	-0.32%
Ligth – tagging \downarrow	$+0.15\%$	$+0.8\%$	$+0.6\%$
Electron ident. \uparrow	-2.4%	–	-1.0%
Electron ident. \downarrow	$+2.4\%$	–	$+1.0\%$
Electron rec. \uparrow	-1.90%	–	-0.84%
Electron rec. \downarrow	$+2.0\%$	–	$+0.85\%$
Electron smear. \uparrow	$+0.13\%$	–	$+0.06\%$
Electron smear. \downarrow	$+0.4\%$	–	$+0.16\%$
Electron scale. \uparrow	$+0.4\%$	–	$+0.16\%$
Electron scale. \downarrow	-0.44%	–	-0.17%
Muon MS \uparrow	–	-0.86%	-0.06%
Muon MS \downarrow	–	-0.09%	-0.04%
Muon ID \uparrow	–	$+0.04\%$	$+0.02\%$
Muon ID \downarrow	–	$+0.02\%$	$+0.01\%$
Muon scale	–	$\pm 0.20\%$	$\pm 0.11\%$
Muon ident. \uparrow	–	-0.76%	-0.43%
Muon ident. \downarrow	–	$+0.78\%$	$+0.43\%$
MPI \uparrow	-0.42%	$+0.66\%$	$+0.18\%$
MPI \downarrow	$+0.21\%$	-0.33%	-0.1%
HFOR	$\pm 3.7\%$	$\pm 3.0\%$	$\pm 3.2\%$
Trigger \uparrow	-0.06%	-0.54%	-0.34%
Trigger \downarrow	$+0.06\%$	$+0.57\%$	$+0.35\%$

Table 7.2: Systematic uncertainty components for the electron, muon and combined channel.

7.4 Results

The n_{bb} signal event yield extracted from the data in the electron, muon and combined channels via the template-fit method are given in Table 7.3 along with the corresponding statistical uncertainties.

Figure 7.3 shows the pre-fit (first column) and post-fit (second column) template distributions in the electron (top row), muon (middle row) and combined (bottom row) channel.

The cross-section estimates for the production of a Z boson candidate in association with a bottom-antibottom pair, in the 2011 data samples used, in the electron and muon channel are:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow e^+e^-) = 436.3 \pm 32.4 \text{ (stat.) } {}^{+52.3}_{-50.3} \text{ (syst.) } \pm 7.9 \text{ (lumi.) fb, (7.8)}$$

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow \mu^+\mu^-) = 552.0 \pm 33.6 \text{ (stat.) } {}^{+64.4}_{-59.6} \text{ (syst.) } \pm 10.0 \text{ (lumi.) fb. (7.9)}$$

The combined template distributions before (after) performing the template fit are shown in Figure 7.3 left (right) side. The corresponding $Z + b\bar{b}$ cross-section measurement for the combination of the electron and muon channels is:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 499.1 \pm 23.5 \text{ (stat.) } {}^{+57.7}_{-53.2} \text{ (syst.) } \pm 8.9 \text{ (lumi.) fb. (7.10)}$$

A 2σ difference between the electron and muon channel cross-section measurements is found, assuming a 100% correlation between the systematic uncertainties in two channels. A number of detailed studies on the electron-muon channel cross-section difference is presented in Appendix B.

	template	pre-fit yield	post-fit yield
<i>e</i> -ch.	<i>bb</i>	346.5 ± 7.0	359.8 ± 26.7
	<i>nonbb</i>	265.1 ± 8.6	272.4 ± 24.0
	<i>other</i>	59.3 ± 2.5	59.3 ± 2.5
μ -ch.	<i>bb</i>	432.9 ± 7.7	563.0 ± 34.3
	<i>nonbb</i>	354.6 ± 10.2	371.8 ± 30.4
	<i>other</i>	76.2 ± 3.0	76.2 ± 3.0
<i>cb</i> -ch.	<i>bb</i>	779.5 ± 10.4	921.0 ± 43.3
	<i>nonbb</i>	619.7 ± 13.4	646.3 ± 38.6
	<i>other</i>	135.5 ± 3.9	135.5 ± 3.9

Table 7.3: Templates event yields before (third column) and after (fourth column) performing the template fit. The statistical uncertainties are given.

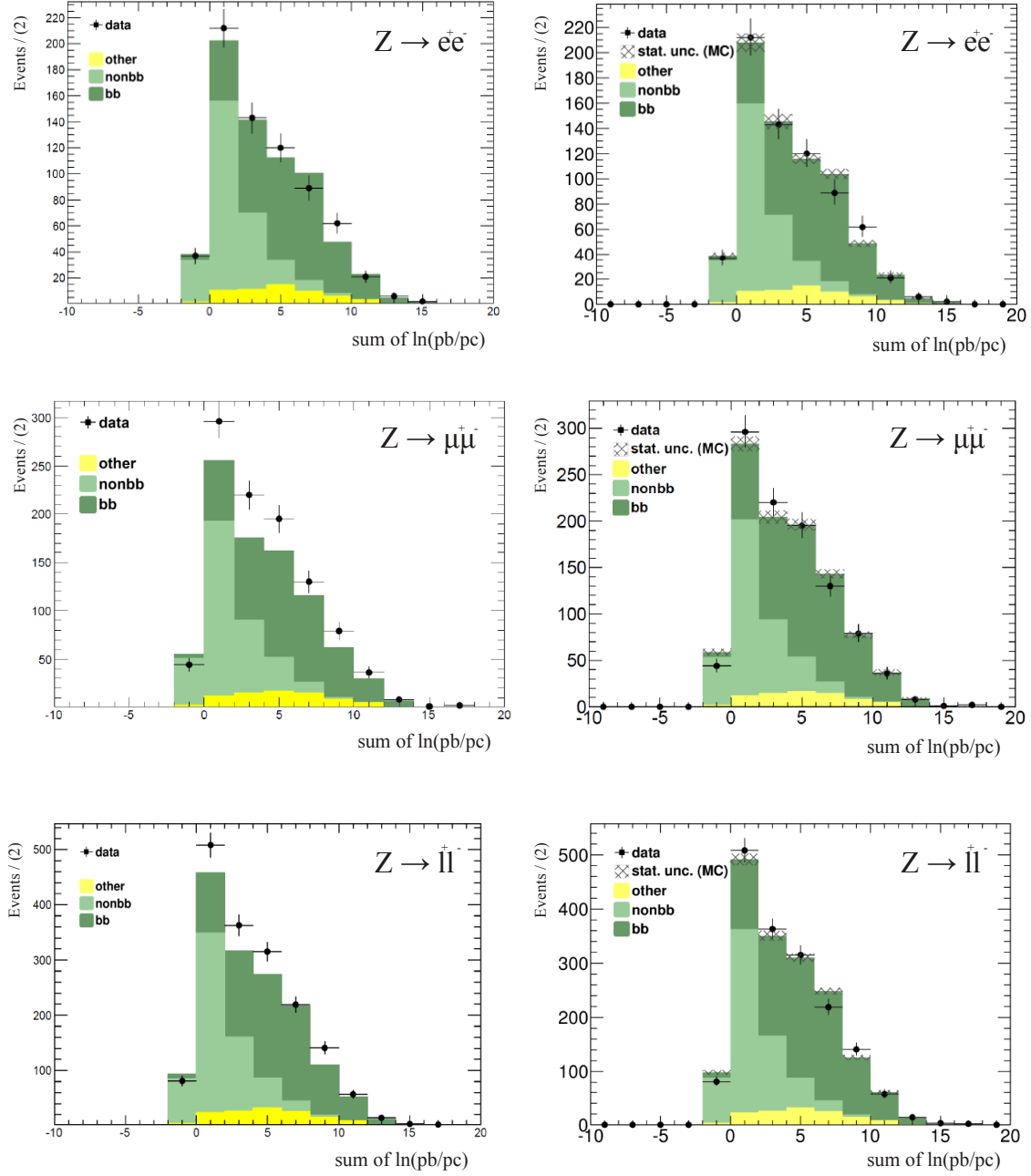


Figure 7.3: Data (points) and simulation (stacked histograms) template distributions in the electron (top row), muon (second row) and combination (bottom row) channel, prior (first column) and subsequent (second column) to the fit to data.

7.5 Theoretical Predictions

A theoretical $Z + b\bar{b}$ cross-section prediction at $\sqrt{s} = 7$ TeV [66] has been evaluated with MCFM [59] ³ following closely the signal event selection described in Sections 5.1 and 5.2. All leptons were required to have $p_T > 20$ GeV, $|\eta| < 2.5$ and an invariant mass within 76 GeV - 106 GeV. Jets, with $p_T > 20$ GeV and $|y| < 2.4$, reconstructed from *parton* clusters, were selected. *B*-jets have been defined at parton level using the same ΔR matching procedure as described in Section 7.2 with the difference that *b*-hadrons were replaced with *b*-quarks in the matching.

In order to account for the absence of QED final state radiation effects in MCFM, a correction has been estimated with the Alpgen generator by taking the signal yield ratio using the same *particle level* selection as described in Section 7.2 requiring *dressed* leptons (absent in MCFM) in one case and *Born* leptons (present in MCFM) in the other.

Non-perturbative QCD corrections ⁴, were estimated using a set of signal samples simulated with Sherpa 1.4.1, allowing for a convenient shutdown of the fragmentation and underlying event. The non-perturbative QCD correction was estimated by taking the ratio of the signal yield in events with a Z boson and two jets matched with either *b*-hadrons or with *b*-quarks using the default ΔR matching procedure (see Section 7.2). For the evaluation of a systematic uncertainty on the non-perturbative correction, the Sherpa samples were replaced with PYTHIA 6.427 samples and the signal event selection was repeated. The differences with respect to the central MCFM prediction from Sherpa and Pythia were averaged and taken as a non-perturbative systematic uncertainty.

The renormalization and factorization scales have been set on event-by-event basis to the quadrature sum of the Z boson invariant mass and transverse momentum, i.e.:

$$Q^2 = m_Z^2 + p_{T,Z}^2 \quad (7.11)$$

A systematic uncertainty on the chosen scales has been estimated by shifting their values up or down by a factor of 2 independently.

A second source of theoretical systematic uncertainty, related to the strong coupling constant α_s , was estimated by shifting its default value by $\pm 1\sigma$ in two independent steps and taking the difference with respect to the nominal result in each case as an uncertainty.

Finally, the systematic uncertainty related to the MCFM input CTEQ6.6 PDF was estimated by adding in quadrature all the positive (negative) variations of each of the twenty-two eigenvectors [62].

The theoretical $Z + b\bar{b}$ MCFM cross-section prediction was estimated to be:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+ l^-) = 403_{-66}^{+82} \text{ fb} \quad (7.12)$$

³MCFM is a next-to-leading order MonteCarlo program-tool with a parton level matrix element framework, designed for computing a large spectrum of hadron collider physical processes cross-sections.

⁴Such as those related to the hadronization and the underlying event absent in the MCFM framework.

8 Summary and Conclusions

A first cross-section measurement of $Z + b\bar{b}$ production with the ATLAS detector using the full dataset of $\int L(t) dt = 4.64 \text{ fb}^{-1}$ of $\sqrt{s} = 7 \text{ TeV}$ data collected in the year 2011 is presented.

A $Z + b\bar{b}$ cross-section measurement provides a direct test of next-to-leading order perturbative QCD predictions. Secondly, it provides an estimate of the highest irreducible background for the associate production of the electroweak symmetry breaking scalar boson with a Z boson as well as for various other processes beyond the Standard Model.

The Z boson decay channels considered are: $Z \rightarrow e^+e^-$ and $Z \rightarrow \mu^+\mu^-$.

Exactly two opposite-charged same-flavor leptons with $p_T^{\text{lepton}} > 20 \text{ GeV}$, $|\eta^{\text{lepton}}| < 2.4$ and invariant mass window of $76 \text{ GeV} < M_{ll} < 106 \text{ GeV}$ produced in association with at least two b -tagged jets with $p_T^{\text{jet}} > 20 \text{ GeV}$ and $|y^{\text{jet}}| < 2.4$ are selected in the signal region.

A likelihood template-fit method seeded by simulation-based signal and background templates is used for the signal event yield extraction from the data sample.

The measured $Z + b\bar{b}$ event yield prediction in the electron (muon) channel is: 360 (563) events with a background estimate of 332 (448) events. The corresponding electron and muon channel cross-section measurement results are:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow e^+e^-) = 436.3 \pm 32.4 \text{ (stat.) } {}^{+52.3}_{-50.3} \text{ (syst.) } \pm 7.9 \text{ (lumi.) fb,} \quad (8.1)$$

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow \mu^+\mu^-) = 552.0 \pm 33.6 \text{ (stat.) } {}^{+64.4}_{-59.6} \text{ (syst.) } \pm 10.0 \text{ (lumi.) fb.} \quad (8.2)$$

and:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 499.1 \pm 23.5 \text{ (stat.) } {}^{+57.7}_{-53.2} \text{ (syst.) } \pm 8.9 \text{ (lumi.) fb,} \quad (8.3)$$

for the combined channel. A fair agreement with a NLO theoretical prediction is found:

$$\sigma(Z + b\bar{b}) \cdot Br(Z \rightarrow l^+l^-) = 403 {}^{+82}_{-66} \text{ fb.} \quad (8.4)$$

The main source of systematic uncertainty, on the order of 10%, is caused by the uncertainty on the b -tagging of jets.

A difference of approximately two standard deviations between the electron and muon channel cross-section measurements, caused by either an excess of data events in the muon channel or a deficit of data events in the electron channel is observed. The source of the discrepancy has been investigated in terms of event selection stages, Z -boson candidate data excess as a function of trigger and pileup conditions, b -tagging systematic effects, isolated muons originating in semileptonic decays in heavy-jets and various object kinematic distributions.

It has been found that a large fraction of the discrepancy is caused by an excess of data events in the muon channel in the region within 20 GeV–30 GeV of the sub-leading b -tagged jet transverse momentum distribution. Additional object kinematic distributions for events in this region have not revealed any obvious sources of the difference.

A Appendix

This section describes the details of the multijet statistical uncertainty evaluation in the electron and muon channels, as referred to in Chapter 6. The method uses the error propagation formula of a linear function with n parameters — $f(x_1, x_2, \dots, x_n)$:

$$\sigma f(x_1, x_2, \dots, x_n) = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right) \cdot \left(\frac{\partial f}{\partial x_j} \right) \sigma_{x_i} \sigma_{x_j} \cdot \rho_{ij}}, \quad (\text{A.1})$$

where σ_{x_i} and $\sigma f(x_1, x_2, \dots, x_n)$ are the uncertainties of the x_i parameter and that of the function respectively and ρ_{ij} is the correlation coefficient between parameters x_i and x_j . The above expression can be further split into:

$$\sigma f(x_1, x_2, \dots, x_n) = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_{x_i}^2 + 2 \sum_{\substack{i,j=1 \\ i < j}}^n \left(\frac{\partial f}{\partial x_i} \right) \cdot \left(\frac{\partial f}{\partial x_j} \right) \sigma_{x_i} \sigma_{x_j} \cdot \rho_{ij}}. \quad (\text{A.2})$$

Electron Channel

Table A.1 shows the data and simulation event yields in the opposite (OS) and same charge sign (SS) channels, inside the *Z-peak Region* (76 GeV – 106 GeV) and *Control Region* (51 GeV – 131 GeV), corresponding to the invariant mass distributions from Figures 6.4 and 6.5 but before applying the *Z-peak Region* normalization scale factors to simulation.

Mass Window	$D_{\text{BJS}}^{\text{OS}}$	$S_{\text{BJS}}^{\text{OS}}$	$D_{\text{BJS}}^{\text{SS}}$	$S_{\text{BJS}}^{\text{SS}}$	$D_{\text{AJS}}^{\text{SS}}$	$S_{\text{AJS}}^{\text{SS}}$
<i>Z-peak Region</i>	1008769	965406	8251	10991	5	7.82
<i>Control Region</i>	1083292	1031850	9884	12024	12	10.47

Table A.1: Data and simulation event yields in the electron channel, before applying the *Z-peak Region* normalization scale factors to simulation.

- $D_{\text{BJS}}^{\text{OS}}$ ($S_{\text{BJS}}^{\text{OS}}$) represents the number of data (simulation) events from *Region B* (before jet selection, opposite charge sign).
- $D_{\text{BJS}}^{\text{SS}}$ ($S_{\text{BJS}}^{\text{SS}}$) represents the number of data (simulation) events from *Region D* (before jet selection, same charge sign).
- $D_{\text{AJS}}^{\text{SS}}$ ($S_{\text{AJS}}^{\text{SS}}$) represents the number of data (simulation) events from *Region C* (after jet selection, opposite charge sign).

The multijet background event yields given in Table 6.1 were estimated as follows:

$$\text{QCD}^{\text{B}} = {}^{\text{cr}}D_{\text{BJS}}^{\text{OS}} - \frac{{}^{\text{z}}D_{\text{BJS}}^{\text{OS}}}{{}^{\text{z}}S_{\text{BJS}}^{\text{OS}}} \cdot {}^{\text{cr}}S_{\text{BJS}}^{\text{OS}} \quad (\text{A.3})$$

$$\text{QCD}^{\text{C}} = ({}^{\text{cr}}D_{\text{AJS}}^{\text{SS}} - {}^{\text{z}}D_{\text{AJS}}^{\text{SS}}) - \frac{{}^{\text{z}}D_{\text{BJS}}^{\text{SS}}}{{}^{\text{z}}S_{\text{BJS}}^{\text{SS}}} \cdot ({}^{\text{cr}}S_{\text{AJS}}^{\text{SS}} - {}^{\text{z}}S_{\text{AJS}}^{\text{SS}}) \quad (\text{A.4})$$

$$\text{QCD}^{\text{D}} = {}^{\text{cr}}D_{\text{BJS}}^{\text{SS}} - \frac{{}^{\text{z}}D_{\text{BJS}}^{\text{SS}}}{{}^{\text{z}}S_{\text{BJS}}^{\text{SS}}} \cdot {}^{\text{cr}}S_{\text{BJS}}^{\text{SS}} \quad (\text{A.5})$$

$$QCD^A = QCD^C \cdot \frac{QCD^B}{QCD^D} \quad (A.6)$$

where the “cr” and “z” superscripts have been applied in order to emphasise the event yields measured inside the *Control Region* or inside the *Z-peak Region* windows respectively ¹.

The statistical uncertainty of each of the multijets components from the four fiducial regions has been evaluated using relation (A.2). The statistical uncertainties on the data coefficients from the above relations were taken as square root of the corresponding event yields while the same uncertainties in simulation were taken as the square root of the quadrature sum of the statistical uncertainties of each contributing background channel weighted by luminosity.

Muon Channel

In the muon channel the same method for the evaluation of the multijet statistical uncertainties in the four fiducial regions as in the electron channel was used. Table A.2 gives the data and simulation event yields in the isolated (iso) and anti-isolated (anti-iso) channels, inside the *Z-peak Region* and *Control Region*, corresponding to the invariant mass distributions from Figures 6.8 and 6.9 before applying the *Z-peak Region* normalization scale factors to simulation.

Mass Window	D_{BJS}^{iso}	S_{BJS}^{iso}	$D_{BJS}^{anti-iso}$	$S_{BJS}^{anti-iso}$	$D_{AJS}^{anti-iso}$	$S_{AJS}^{anti-iso}$
<i>Z-peak Region</i>	1395551	1354330	2198	61.8	6	0.26
<i>Control Region</i>	1509761	1459360	7639	103.6	21	0.67

Table A.2: Data and simulation event yields in the muon channel, before applying the *Z-peak Region* normalization scale factors to simulation.

- D_{BJS}^{iso} (S_{BJS}^{iso}) represents the number of data (simulation) events from *Region B* (before jet selection, isolated muons).

¹For example $^{cr}D_{BJS}^{OS}$ is the number of data events, inside the *Control Region* from *Region B* and $^zS_{BJS}^{SS}$ corresponds to the number of simulation events inside the *Z-peak Region* from *Region D*.

- $D_{\text{BJS}}^{\text{anti-iso}}$ ($S_{\text{BJS}}^{\text{anti-iso}}$) represents the number of data (simulation) events from *Region D* (before jet selection, anti-isolated muons).
- $D_{\text{AJS}}^{\text{anti-iso}}$ ($S_{\text{AJS}}^{\text{anti-iso}}$) represents the number of data (simulation) events from *Region C* (after jet selection, isolated muons).

The multijet background contributions shown in table 6.2 were evaluated with the following set of relations:

$$\text{QCD}^{\text{B}} = \text{cr}D_{\text{BJS}}^{\text{iso}} - \frac{zD_{\text{BJS}}^{\text{iso}}}{zS_{\text{BJS}}^{\text{iso}}} \cdot \text{cr}S_{\text{BJS}}^{\text{iso}} \quad (\text{A.7})$$

$$\text{QCD}^{\text{C}} \approx \text{cr}D_{\text{AJS}}^{\text{anti-iso}} - zD_{\text{AJS}}^{\text{anti-iso}} \quad (\text{A.8})$$

$$\text{QCD}^{\text{D}} \approx \text{cr}D_{\text{BJS}}^{\text{anti-iso}} - zD_{\text{BJS}}^{\text{anti-iso}} \quad (\text{A.9})$$

$$\text{QCD}^{\text{A}} = \text{QCD}^{\text{C}} \cdot \frac{\text{QCD}^{\text{B}}}{\text{QCD}^{\text{D}}} \quad (\text{A.10})$$

B Appendix

In the following appendix a detailed summary of the studies made on the electron-muon channel discrepancy will be presented.

B.1. Event Selection Stages

Z + at least two jets

Figure B.1 shows the leading (left) and sub-leading (right) jet MV1 weight distribution in the muon (top two plots) and electron (bottom two plots) channels, corresponding to an event selection with a Z boson candidate and at least two jets prior to the b -tagging selection. Table B.1 gives the data and simulation event yields. An overestimation of the ALPGEN prediction for the light-jet contribution at low MV1 jet weight values as well as an underestimation of the heavy-jet event rate at high MV1 weight values are observed. However, the overall agreement in terms of total event yields between data and simulation in the two channels was found to be satisfactory.

Z + at least one b -jet

By requiring at least one b -tagged jet in the event, the relative excess in data was found to be at the same 14% level in both lepton channels. The leading (in p_T) MV1 b -tagged jet-weight distribution in the muon (third row, left) and electron (fourth row, left) channels are shown in Figure B.1. Additionally the inclusive MV1 jet-weight distributions for all the other jets in the event in the muon (third row, right) and electron (fourth row, right) channels are shown in the same figure. The corresponding event/jet yields in each channel are given in table B.2. A good agreement between simulation and data at this selection point is not to be expected since the multijets background as well as the $Z + bc$ and $Z + bl$ fraction corrections extracted via the template fit [66] were not included.

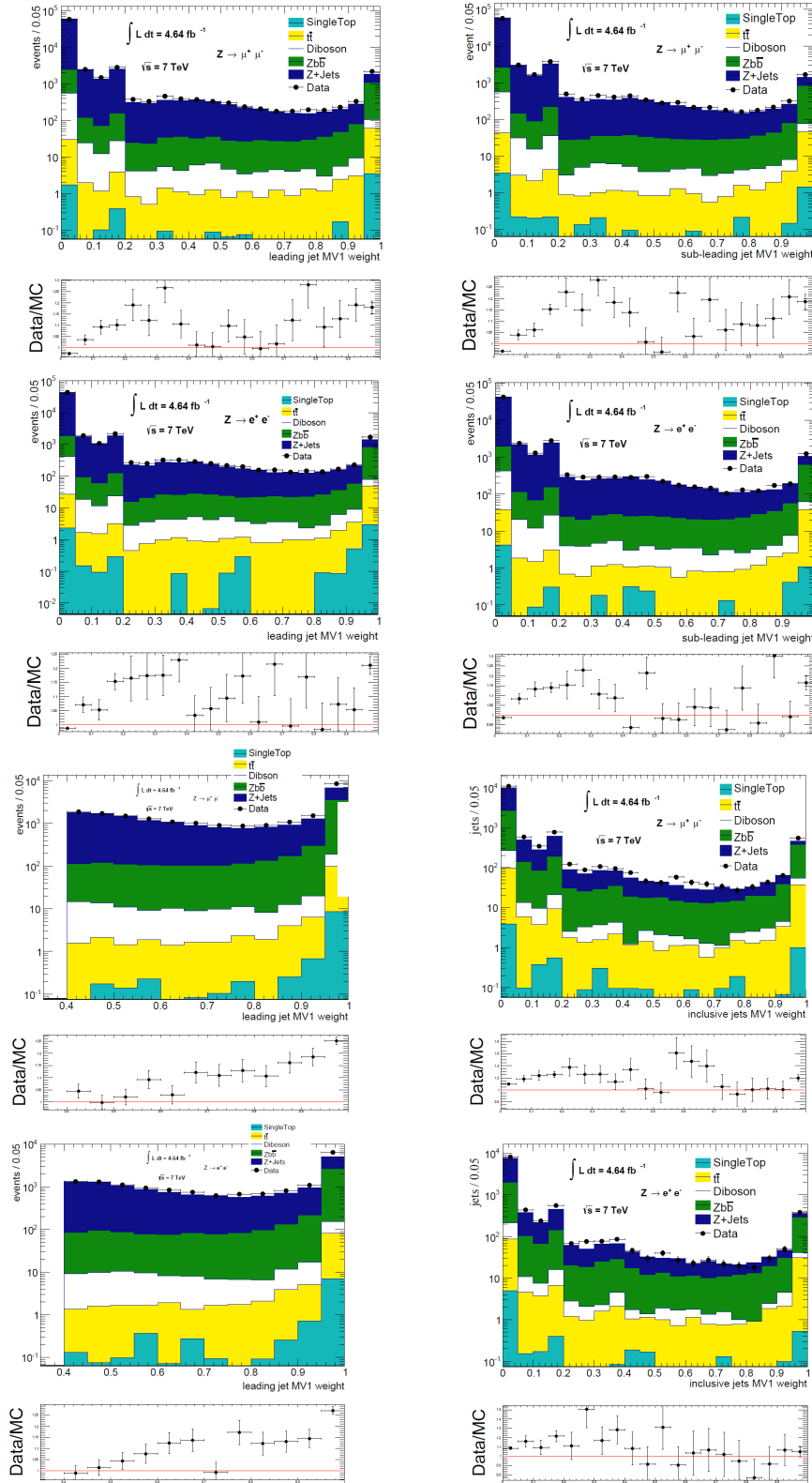


Figure B.1: MV1 weight distribution of the leading jet (first and second row, left) and sub-leading jet (first and second row, right) for an event selection with a Z boson and at least two jets. The third and fourth rows, left, show the leading MV1 b-tagged jet weight distributions; the third and fourth rows, right, show the MV1 weight distributions for all other jets (i.e. excluding the leading b-tagged jet) for an event selection with a Z boson and at least one b-tagged jet. First and third (second and fourth) row distributions correspond to the muon (electron) channel.

	e-channel (event yield)	μ -channel (event yield)
SingleTop	7 ± 0.8	6.1 ± 0.8
$t\bar{t}$	93 ± 2.4	112.1 ± 2.7
Diboson	509.2 ± 9.04	673.6 ± 10.2
$Zb\bar{b}$	2758 ± 17.2	3668 ± 20.0
Z + Jets	49249 ± 105.7	67443 ± 122.9
MC Total	52616 ± 107.5	71902.8 ± 125
Data	53149 ± 230.5	71515 ± 267.4
Relative excess	$1\% \pm 0.48\%$	$-0.54\% \pm 0.41\%$

Table B.1: Data and simulation event yields and their corresponding statistical uncertainties in the electron (second column) and muon (third column) channels for an event selection with a Z boson candidate in association with at least two jets.

	e-channel (event yield)	e-channel (inclusive non-tagged jets yield)	μ -channel (event yield)	μ -channel (inclusive non-tagged jets yield)
SingleTop	9.2 ± 1	6.8 ± 0.85	10.4 ± 1.0	6.8 ± 0.9
$t\bar{t}$	97 ± 2.5	145 ± 3.1	116.5 ± 2.7	168.7 ± 3.4
Diboson	150.5 ± 5.1	167.5 ± 5.3	193.8 ± 6.0	227.5 ± 6.3
$Zb\bar{b}$	3539.6 ± 20.1	2570 ± 17.0	4730.8 ± 23.5	3431 ± 19.8
Z + Jets	10846.3 ± 70.5	6615 ± 39.3	14539.3 ± 80.8	8753.2 ± 44.9
MC Total	14642 ± 73.5	9504 ± 43.3	19590.9 ± 84.4	12587 ± 49.6
Data	16723 ± 129.3	10406 ± 102.0	22329 ± 149.4	14116 ± 118.8
Relative excess	$14.2\% \pm 1.05\%$	$9.50\% \pm 1.18\%$	$14.0\% \pm 0.9\%$	$12.1\% \pm 1\%$

Table B.2: Data and simulation event yields in the electron (second column) and muon (fourth column) channel corresponding to an event selection with Z boson and at least one b-tagged jet. Inclusive non-tagged jet yields in the electron (third column) and muon (fifth column). The statistical uncertainties are given.

However, the amount of data excess in both the electron and muon channels was found to be of the same order.

Z + at least two b -jets

After the selection of a second b -tagged jet, the data excess in the electron channel decreases to 3.1% while the excess in the muon channel grows to 17%¹. The MV1 jet-weight distributions in the muon (top row) and electron (bottom row) channels are shown in Figure B.2 for the leading (left side) and sub-leading (right side) jets. The data and simulation event yields are given in table B.3. A good overall agreement between data and simulation MV1 jet weight distributions in the electron channel is observed while the muon channel distributions present an excess of events of which a significant fraction is located in the last bin.

	e-channel (event yield)	μ -channel (event yield)
SingleTop	1.1 ± 0.3	1.4 ± 0.4
$t\bar{t}$	41 ± 1.6	48.4 ± 1.9
Diboson	17.2 ± 1.9	26.5 ± 2.3
$Zb\bar{b}$	387 ± 6.9	495.3 ± 7.9
Z + Jets	224.6 ± 8.7	292.2 ± 10.1
MC Total	671 ± 11.4	863.7 ± 13.1
Data	692 ± 26.3	1011 ± 31.8
Relative excess	$3.1\% \pm 4.3\%$	$17.1\% \pm 4.1\%$

Table B.3: Data and simulation event yields in the electron (second column) and muon (third column) channels corresponding to a Z + at least two b -tagged jets event selection level. The statistical uncertainties are given.

¹Such a discrepancy between the electron and muon channels is already suggested by the MV1 inclusive jet weight distributions from Figure B.1 (third and fourth rows, right side) at jet weights above 0.4, particularly in the last bin.

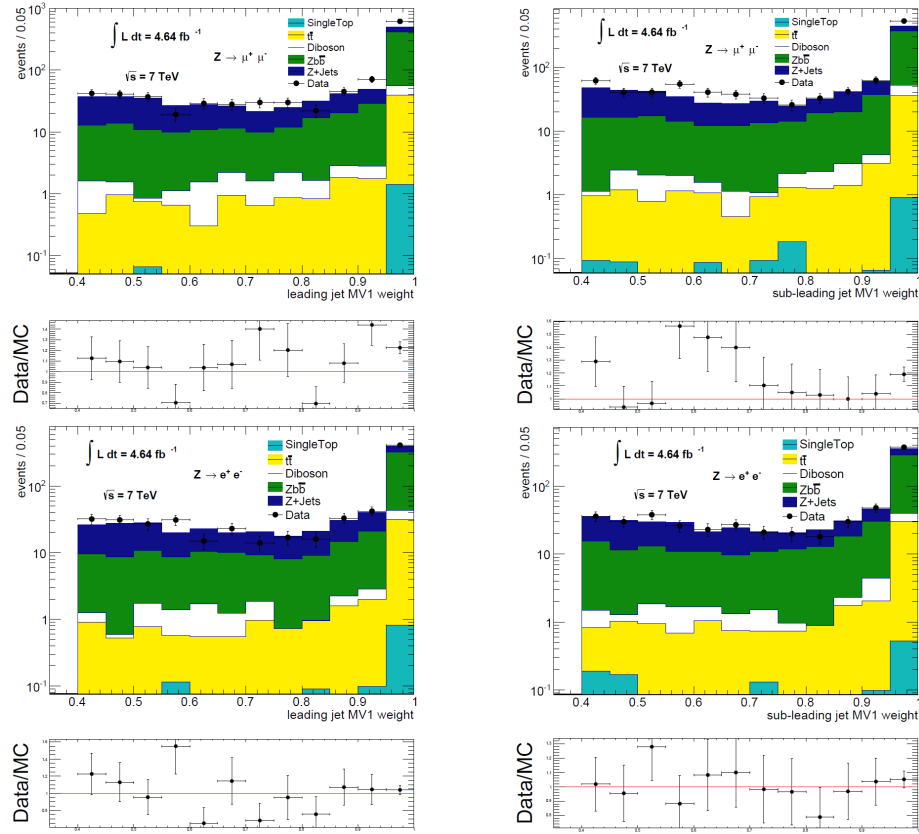


Figure B.2: Leading (first column) and sub-leading (second column) b -tagged jets MV1 weight distribution in the muon (top row) and electron (bottom row) channels for an event selection with a Z boson candidate and at least two b -tagged jets.

B.2. Data Excess Kinematics

Figure B.3, first row, shows the invariant mass distribution of the signal two leptons in the electron (left side) and muon (right side) channels. The second row, shows the missing transverse energy distributions (E_T^{miss}) in the electron/muon channels (left/right side), for events with a Z boson candidate and at least two b -tagged jets.

The four plots shown in the first and second row in Figure B.4 correspond to the data minus simulation MV1 jet-weight distributions while the third and fourth rows show the data excess in the distributions shown in Figure B.3 ². While the electron channel is characterized by a reasonable agreement between simulation and data, the muon channel presents an excess of data events.

The excess found at high MV1 jet-weight values in the muon channel, suggest a high heavy-jet origin likelihood.

The excess centered around the Z boson mass (Figure B.4 third row, right distribution) as well as the bottom row right distribution from the same Figure, corresponding to the missing transverse energy distribution of the excess, strongly discriminate between $t\bar{t}$ -like backgrounds (e.g. $t\bar{t}$, SingleTop, multijets, etc.) and events with a real Z boson such as $Z + b\bar{b}$, Z +light-jets and diboson.

Due to the relative small event yield for diboson processes ³ and the tendency of light-quark induced jets to accumulate at low MV1 jet-weights, the data excess kinematic distributions indicate that the excess observed in the muon channel resembles the signal.

B.3. B -tagging systematic effects

The scenario according to which the electron-muon channel cross-section difference is caused by a faulty b -tagging efficiency/inefficiency correction in simulation has been considered. The pre-fit template distributions decomposed into individual jet flavor ⁴ for the muon (first row) and electron (second row) channels are shown in Figure B.5 ⁵.

²The excess distributions were obtained by subtracting simulation from the data.

³The individual muon channel diboson contribution given in table B.3 is: ZZ ($Z \rightarrow ll$, $Z \rightarrow q\bar{q}$) – 22.9 events; W^+Z ($W^+ \rightarrow q\bar{q}$, $Z \rightarrow ll$) – 2.1 events and W^-Z ($W^- \rightarrow q\bar{q}$, $Z \rightarrow ll$) – 1.5 events. None of the individual diboson contributions as well as their sum could explain the large data excess (around 150 events) observed in the muon channel.

⁴The jet flavor was assessed by matching reconstructed selected jets with particle level quarks with $p_T^{\text{quark}} > 5 \text{ GeV}$ if $\Delta R_{\text{selected-jet, quark}} < 0.5$.

⁵Unlike the templates used for the $Z + b\bar{b}$ cross-section measurement (bb , $nonbb$ and $other$), the components from Figure B.5 were generated from both signal and background samples.

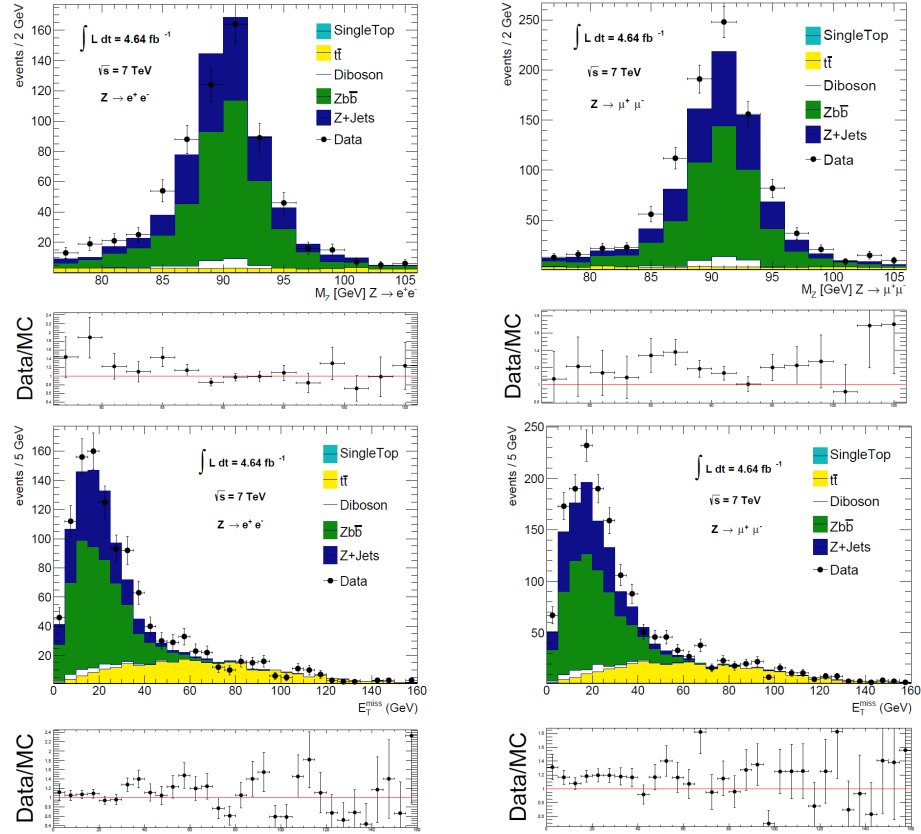


Figure B.3: Data and simulation invariant mass (top row) and missing transverse energy (bottom row) distributions in the electron (left side) and muon channels (right side) corresponding to events with a Z boson candidate and at least two b-tagged jets.

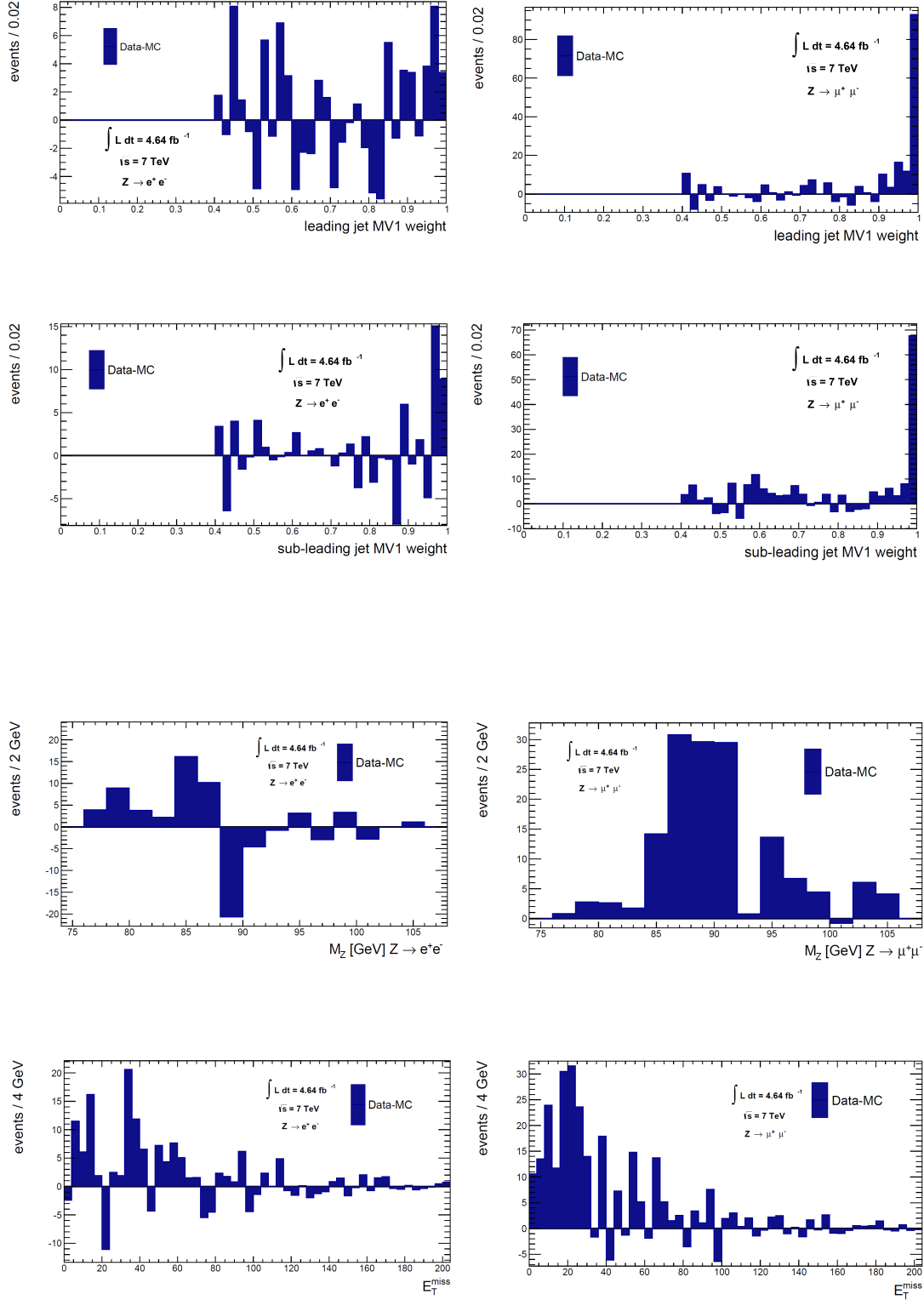


Figure B.4: Data minus simulation MV1 weight distribution of the leading (first row) and sub-leading (second row) jets. Invariant mass (third row) and E_T^{miss} (fourth row) data excess distributions. Distributions on the left (right) side correspond to the electron (muon) channel. All distributions correspond to an event selection with a Z boson and at least two b-tagged jets.

The left side distributions have all b -tagging corrections applied, those on the right side have no efficiency/inefficiency corrections applied. The individual flavor event yields in the electron and muon channels corresponding to events with/without the b -tagging corrections are given in Table B.4. An increase of the data excess in both channels when the b -tagging corrections were not applied is observed. However, the difference between the electron and muon channels still persists.

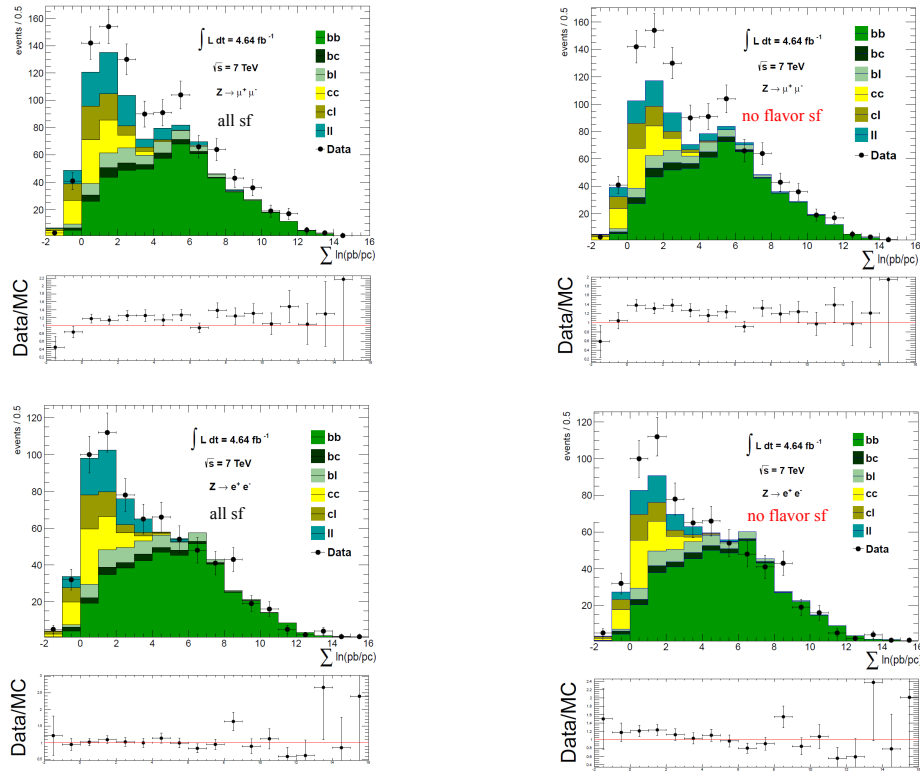


Figure B.5: Pre-fit templates decomposed into individual jet flavor in the muon (first row) and electron (second row) channels for an event selection with a Z boson candidate and at least two b -tagged jets. Left side distributions contain all b -tagging corrections while for those on the right the b -tagging corrections were not included.

Flavor	e-channel with b -corr (event yield)	μ -channel with b -corr (event yield)	e-channel without b -corr (event yield)	μ -channel without b -corr (event yield)
bb	398.3 ± 7.4	499.8 ± 8.2	426.3 ± 7.9	532.9 ± 8.7
bc	23.3 ± 1.7	32.5 ± 2.2	22.1 ± 1.5	31.3 ± 2.1
bl	52.9 ± 3.1	62.8 ± 3.0	43.4 ± 2.5	51.6 ± 2.4
cc	73.8 ± 3.7	90.3 ± 3.8	65.1 ± 3.4	80.3 ± 3.4
cl	48.7 ± 3.6	68.5 ± 4.0	36.5 ± 2.7	50.9 ± 3.0
ll	73.9 ± 6.1	109.9 ± 7.9	47.0 ± 3.8	68.9 ± 4.8
MC Total	671 ± 11.4	863.7 ± 13.1	640.5 ± 10.2	816.1 ± 11.4
Data	692 ± 26.3	1011 ± 31.8	692 ± 26.3	1011 ± 31.8
Relative excess	$3.1\% \pm 4.3\%$	$17.1\% \pm 4.1\%$	$8\% \pm 4.5\%$	$23.9\% \pm 4.3\%$

Table B.4: Data and simulation event yields in the electron (second and fourth rows) and muon (third and fifth) channels corresponding to an event selection with a Z boson candidate and at least two b -tagged jets. Events from the second and third rows contain all corrections while those from the fourth and fifth column do not contain the b -tagging corrections.

B.4. Trigger Period Dependence

The data excess in both the electron and muon channels has been examined as a function of the 2011 trigger-periods. The top three (bottom two) row distributions from Figure B.6 show the data excess (data minus simulation) as a function of the signal electrons (muons) invariant mass. Table B.5 gives the event yields for data and simulation, the various triggers and the corresponding trigger integrated luminosity. A fair agreement between the data and simulation in the three trigger-splitted periods in the electron channel is observed while the muon channel presents a consistent data excess centred around the Z boson mass in both trigger periods.

B.5. Pileup Period Dependence

The amount of pileup varied in 2011 data as shown in Figure B.7. A similar pileup distribution was simulated in each 2011 MC11c sample for an equivalent fraction of events as in data. The data excess has been investigated as a function of the pileup conditions in the electron (Figure B.8, top three distributions) and muon channels (Figure B.8, bottom three distributions) via the leading two leptons invariant mass distributions. Due to the relative small integrated luminosity, the first two pileup periods have been merged. The data and simulation event yields corresponding to the investigated pileup periods as well

Trigger	Period	$\int \mathcal{L}(t) dt$	MC yield	Data yield	Excess
EF_2e12_medium	D to J	1659 pb ⁻¹	251 ± 7.3	242 ± 15.6	-3.6% ± 6.8%
EF_2e12T_medium	K	583 pb ⁻¹	88.5 ± 3.8	103 ± 10.1	16.4% ± 12.5%
EF_2e12Tvh_medium	L to M	2402 pb ⁻¹	331.4 ± 7.9	347 ± 18.6	4.7% ± 6.1%
EF_mu18_MG	D to I	1430 pb ⁻¹	282 ± 7.4	336 ± 18.3	19.1% ± 7.2%
EF_mu18_MG_medium	J to M	3213 pb ⁻¹	581.7 ± 10.8	675 ± 25.9	16.0% ± 4.9%

Table B.5: Data (fifth column) and MonteCarlo (fourth column) event yields in the electron (third and fourth rows) and muon (fifth and sixth rows) channels for the triggers used in 2011. The electron and muon channel triggers are shown in the first column as well as the data period interval (second column) and the corresponding integrated luminosity (third column).

their corresponding integrated luminosity and the relative excess are given in table B.6. A consistent excess of events is observed systematically in all pileup-splitting periods in muon channel around the Z boson mass while the electron channel data is well modeled by simulation.

Period	$\int \mathcal{L}(t) dt$	MC yield	Data yield	Relative excess
e-channel, D to H	1102pb ⁻¹	169.3 ± 6.2	166 ± 12.9	-1.9% ± 8.4%
e-channel, I to K	1140pb ⁻¹	170.2 ± 5.5	179 ± 13.4	5.2% ± 8.6%
e-channel, L to M	2402pb ⁻¹	331.4 ± 7.9	347 ± 18.6	4.7% ± 6.1%
μ -channel, D to H	1102pb ⁻¹	214.3 ± 6.4	251 ± 15.8	17.1% ± 8.2%
μ -channel, I to K	1140pb ⁻¹	231.5 ± 6.8	275 ± 16.6	18.8% ± 8%
μ -channel, L to M	2402pb ⁻¹	418 ± 9.2	485 ± 22	16% ± 5.9%

Table B.6: Data (fourth column) and MonteCarlo (third column) event yields in the electron (second, third and fourth rows) and muon (fifth, sixth and seventh rows) channels corresponding to the 2011 pileup conditions. The first column gives the data period intervals in which the pileup conditions were similar; the second column gives the integrated luminosity for each pileup period interval.

B.6. Z Vertex Correction

The impact of the “z vertex corrector” [63] on simulation in both the electron and muon channel has been investigated. Figure B.9 shows the pre-fit template distributions decomposed into jet-flavor for the muon (first row) and electron (second row) channels before/after (left/right side) applying the correction. The data and simulation yields are given in table B.7. A small increase of the order to 2% in simulation in both channels is observed. However, the electron-muon channel discrepancy still remains.

	e-channel yield	μ -channel yield
SingleTop	1.2 ± 0.36	1.4 ± 0.4
$t\bar{t}$	38.0 ± 1.6	45.1 ± 1.9
Diboson	16.7 ± 2.0	27.3 ± 2.5
$Zb\bar{b}$	392.1 ± 7.3	497.8 ± 8.2
Z + Jets	235.3 ± 9.6	305.7 ± 11.2
MC Total	683.3 ± 12.3	877.1 ± 14.2
Data	692 ± 26.3	1011 ± 31.8
Relative excess	$1.3\% \pm 4.3\%$	$15.3\% \pm 4.1\%$

Table B.7: Data and simulation event yields in the electron (second column) and muon (third column) channels corresponding to an event selection with a Z boson candidate and at least two b -tagged jets. The “z vertex correction” has been applied.

B.7. MV1 jet weight distribution at high values

The MV1 weight distributions of the leading and sub-leading b -tagged jets at values close to unity have been investigated. Figure B.10 shows MV1 weight distributions normalized to an area of one event for the leading (top two plots) and sub-leading jets (bottom two plots) for data and simulation in the electron and muon channels. Up to jet-weight values close to 0.99 all distributions present a monotonous horizontal trend and have been deliberately left aside. At values close to 0.997 a peak of the same magnitude can be observed in data/mc as well as in the el/mu – channels.

B.8. Muon-Jet Overlap Removal

The impact of the signal lepton–selected jet Overlap Removal (OR) has been investigated in the muon channel to reveal the potential differences between simulation and data for isolated muons originating in semi-leptonic decays in heavy jets. The minimum separation in ΔR between signal muons and selected jets is shown in Figure B.8 with (right side) and without (left side) applying the OR. The data and simulation muon channel event yield when the OR was not applied are given in table B.8. The muon channel data excess was found to remain at the same 17% level when the OR was not applied.

	μ -channel (event yield)
SingleTop	1.5 ± 0.4
$t\bar{t}$	51.6 ± 1.9
Diboson	29.2 ± 2.5
$Zb\bar{b}$	534.2 ± 8.2
Z + Jets	313.4 ± 10.4
MC Total	929.8 ± 13.6
Data	1088 ± 33
Relative excess	$17.0\% \pm 3.9\%$

Table B.8: Muon channel data and simulation event yields and their corresponding statistical uncertainties in events with a Z boson candidate and at least two b-tagged jets. The muon-jet Overlap Removal was not applied.

B.9. Low p_T b-tagged jets kinematic distributions

The leading and sub-leading selected jet p_T , y and ΔR distributions as well as the signal leptons $\Delta\phi$ are shown in figures B.12 and B.13 for the muon (left side) and electron (right side) channels. All electron channel simulation kinematic distributions were found to be in good agreement with the data. The muon channel data excess for the leading and sub-leading selected jets do not seem to point to a particular region of the detector.

A significant fraction of the muon channel data excess was found in the low region of the transverse momentum spectrum of the sub-leading jet (Figure B.12 second row, left distribution), particularly in the window between 20 GeV - 30 GeV. Additional kinematic distributions for events with selected jets transverse momenta within 20 GeV - 30 GeV are shown in Figure B.14 in terms of η and ϕ angular distributions (first and second rows), flavor fractions (third row), signal muons invariant mass (fourth row, left) and invariant mass excess (fourth row, right). None of the investigated distributions suggested any particular faulty detector region or the presence of an additional omitted background though no final inference can be made given the low event statistics. The signal and background event yields as well those of the templates splitted into jet flavors are given in tables B.10 and B.11 respectively.

B.10. B -taggers and Operating Points

The electron and muon channel cross-sections have been estimated for the different calibrated b -tagging efficiencies (“Operating Points” or OPs) of the MV1 tagger as well as for different b -tagging algorithms and their corresponding Operating Points. The electron and muon channel cross-section results for the used taggers and OPs together with the corresponding statistical uncertainties and the relative cross-section difference are given in table B.9. A variation in terms of measured cross-section values with respect to the Operating Point for the MV1 and JetFitterCOMBNN taggers is observed, particularly between the high and low values of the OPs. A possible cause for such a variation could be caused by the multijets background component which, except for the nominal MV1 at 75% (found to be negligible), was not been estimated. The cross-section difference for each b -tagging algorithm and Operating Point between the electron and muon channels still remains.

B -Tagger	OP	$\sigma^{e-ch.}$ (pb)	$\sigma^{\mu-ch.}$ (pb)	rel. diff. %
MV1	60%	0.4978 ± 0.0396	0.5588 ± 0.0435	12.3 ± 12.5
MV1	70%	0.4352 ± 0.0346	0.5235 ± 0.0360	20.3 ± 12.6
MV1	75%	0.4363 ± 0.0324	0.5520 ± 0.0336	26.5 ± 12.1
MV1	85%	0.4406 ± 0.0313	0.5257 ± 0.0302	19.3 ± 10.9
JetFitterCOMBNN	57%	0.4787 ± 0.0405	0.5395 ± 0.0422	12.7 ± 12.9
JetFitterCOMBNN	60%	0.4639 ± 0.0379	0.5354 ± 0.0407	15.4 ± 12.9
JetFitterCOMBNN	70%	0.4294 ± 0.0336	0.5048 ± 0.0349	17.6 ± 12.3
JetFitterCOMBNN	80%	0.4460 ± 0.0325	0.5423 ± 0.0318	21.6 ± 11.4

Table B.9: Cross-section estimates as function of b -taggers (first column) and Operating Points (second column) for the electron (third column) and muon (fourth column). The fifth column corresponds to the relative cross-section difference defined as $\sigma(Z + b\bar{b}) \cdot \Gamma(Z \rightarrow \mu^+ \mu^-) - \sigma(Z + b\bar{b}) \cdot \Gamma(Z \rightarrow e^+ e^-) / \sigma(Z + b\bar{b}) \cdot \Gamma(Z \rightarrow \mu^+ \mu^-)$.

B.11 Conclusions

The measured cross-section difference between the electron and muon channels has been investigated in terms of event selection stages, b -tagging systematic effects, trigger and pileup dependence, MV1 jet-weight distributions at values close to unity. No obvious source of the difference was found.

Further investigations such the impact of the “*z vertex reweighting*”, muon-jet OR, muon channel kinematic distributions for events with $20 \text{ GeV} < p_T^{\text{selected jet}} < 30 \text{ GeV}$, the use of various taggers and operating points did not bring any new insight on the nature of the difference between the two channels. Finally it should be mentioned that no difference between the electron and muon channels at *particle level* has been observed.

	μ -channel yield
SingleTop	$0. \pm 0.$
$t\bar{t}$	0.57 ± 0.17
Diboson	1.3 ± 0.46
$Zb\bar{b}$	44.5 ± 2.3
$Z + \text{Jets}$	51.1 ± 5.5
MC Total	97.5 ± 6.0
Data	127 ± 11.3
Relative excess	$30.3\% \pm 14.1\%$

Table B.10: Data and simulation muon channel event yields for events with selected jet transverse momenta within 20 GeV - 30 GeV.

Flavor	μ -channel yield
bb	39.7 ± 2.5
bc	3.5 ± 0.9
bl	9.2 ± 1.25
cc	7.0 ± 1.2
cl	8.1 ± 1.6
ll	30.0 ± 4.8
MC Total	97.5 ± 6.0
Data	127 ± 11.3
Relative excess	$30.3\% \pm 14.1\%$

Table B.11: Muon channel pre-fit template distribution event yields splitted into individual jet flavors for events with selected jet transverse momenta within 20 GeV - 30 GeV.

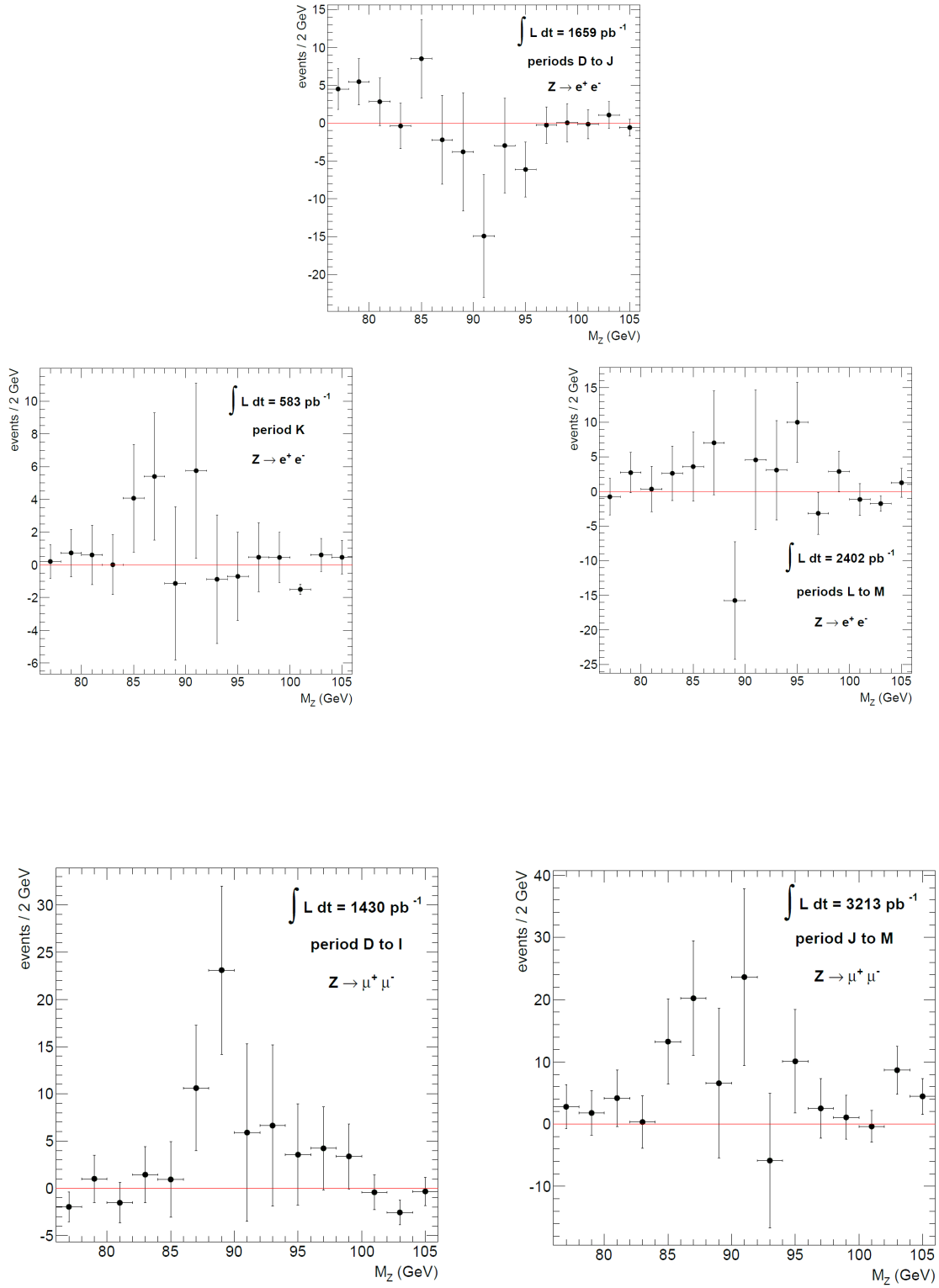


Figure B.6: Signal leptons invariant mass distributions of the data excess in the electron (first row rows) and muon (bottom row) channels splitted according to the 2011 data trigger dependence. All distributions correspond to an event selection with a Z boson candidate and at least two b-tagged jets. The statistical uncertainties are shown.

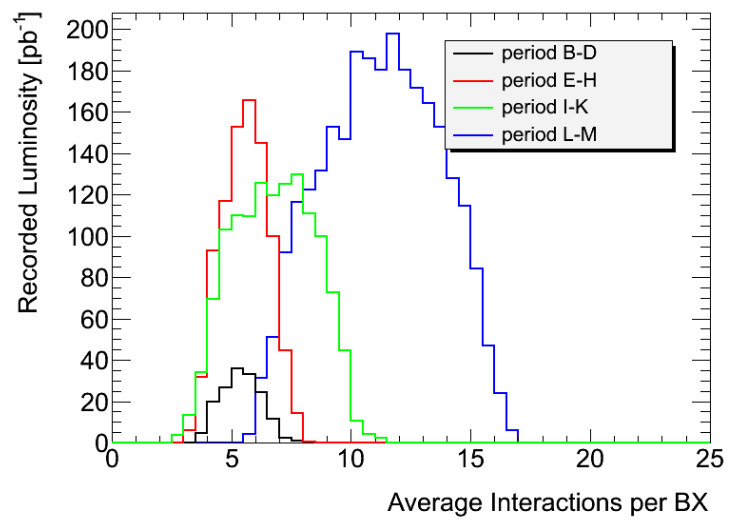


Figure B.7: Pileup period dependence in 2011 data.

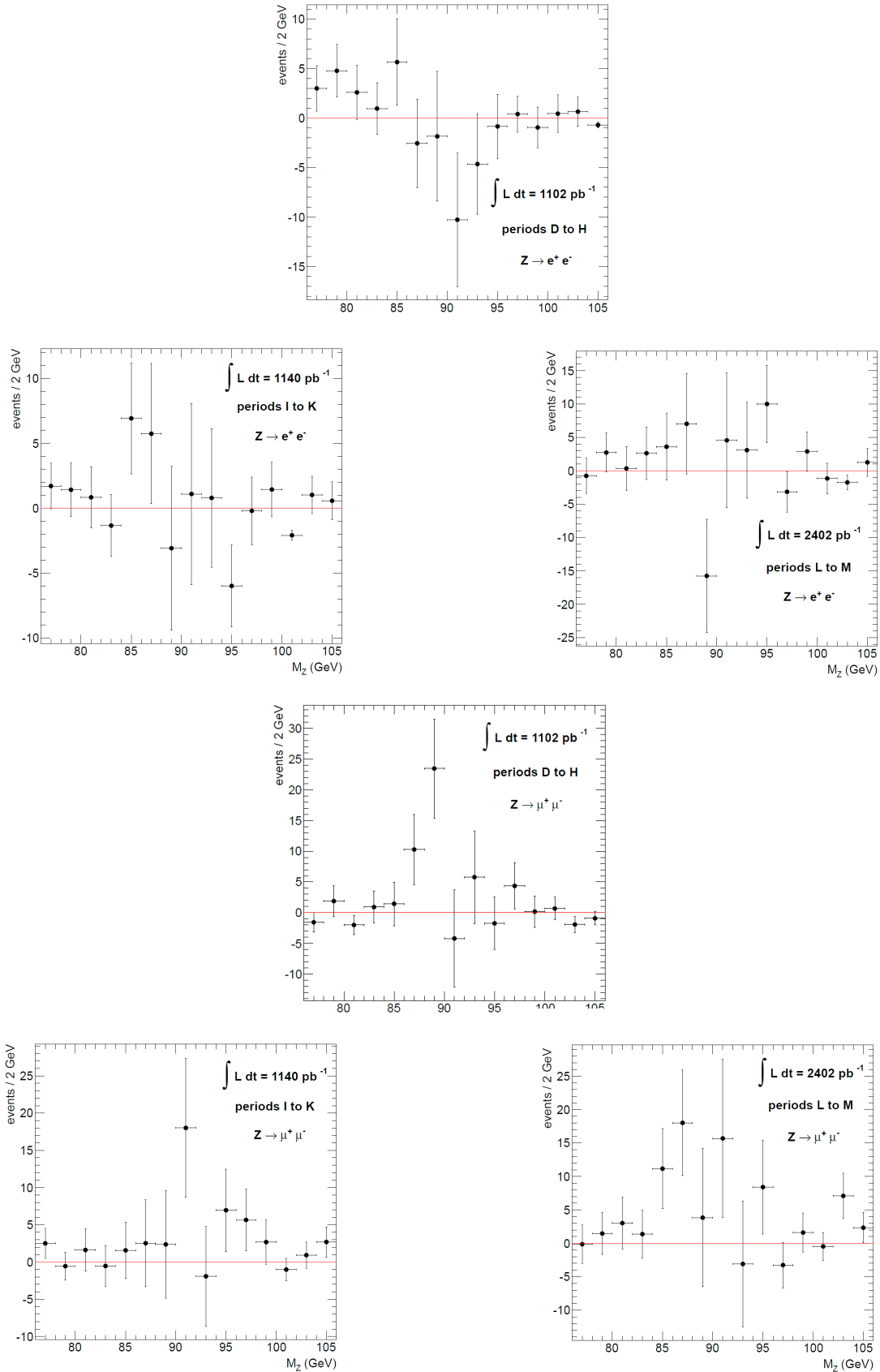


Figure B.8: Signal leptons data excess invariant distributions in the electron (first two rows) and muon (bottom two rows) channels splitted according to the 2011 data pileup dependence. All distributions correspond to an event selection with a Z boson candidate and at least two b-tagged jets. The statistical uncertainties are shown.

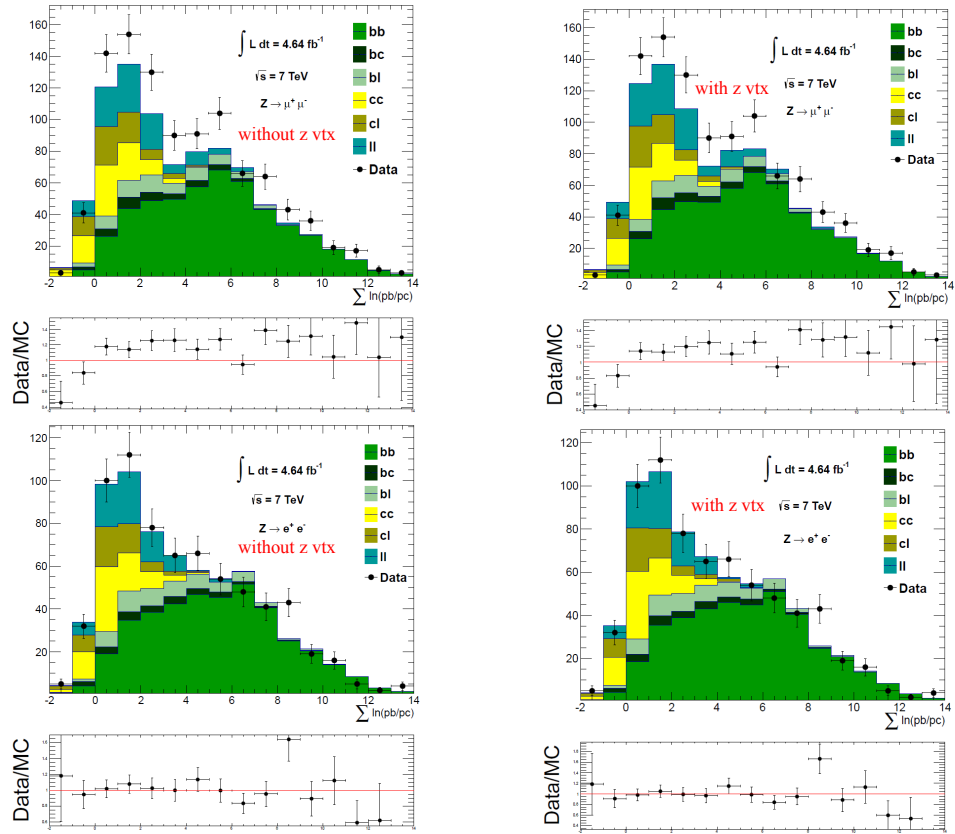


Figure B.9: Pre-fit templates decomposed into individual jet flavor in the muon (top row) and electron (bottom row) channels for an event selection with a Z boson candidate and at least two b-tagged jets. Left side distributions contain all corrections while for those on the right the “z vertex correcton” corrections were not included.

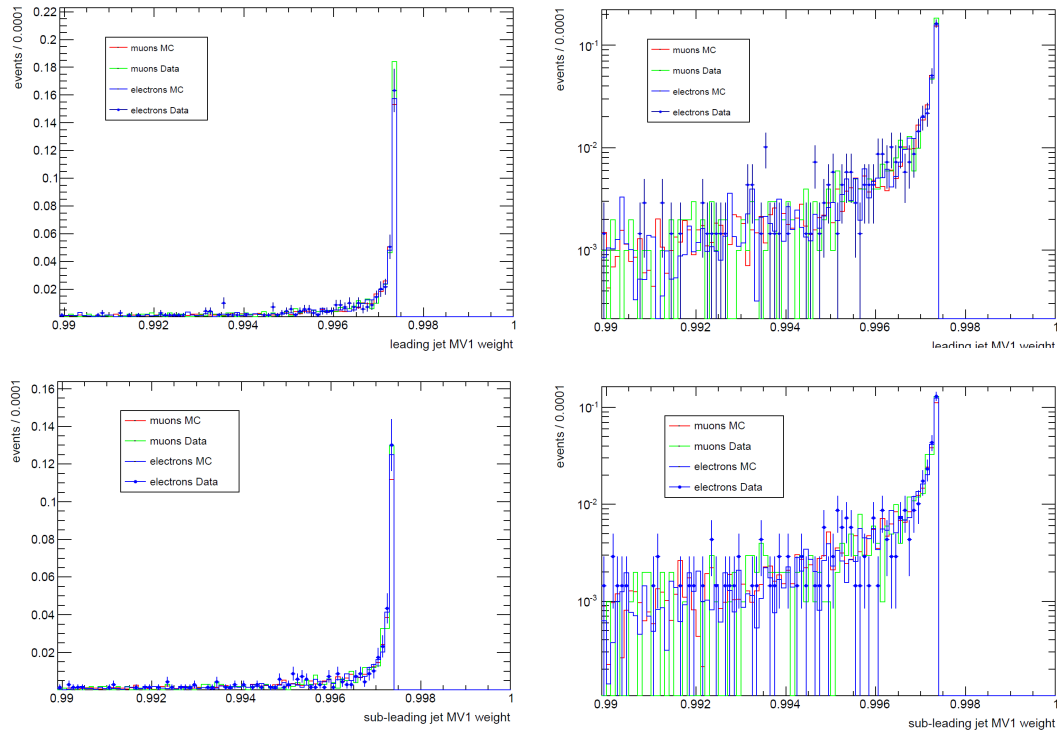


Figure B.10: *MV1 leading (top row) and sub-leading (bottom row) jet weight distributions normalized to unity in the muon / electron channel simulation (red / blue curve) and muon / electron channel data (green curve / blue dots). The first (second) column distributions correspond to a linear (logarithmic) scale.*

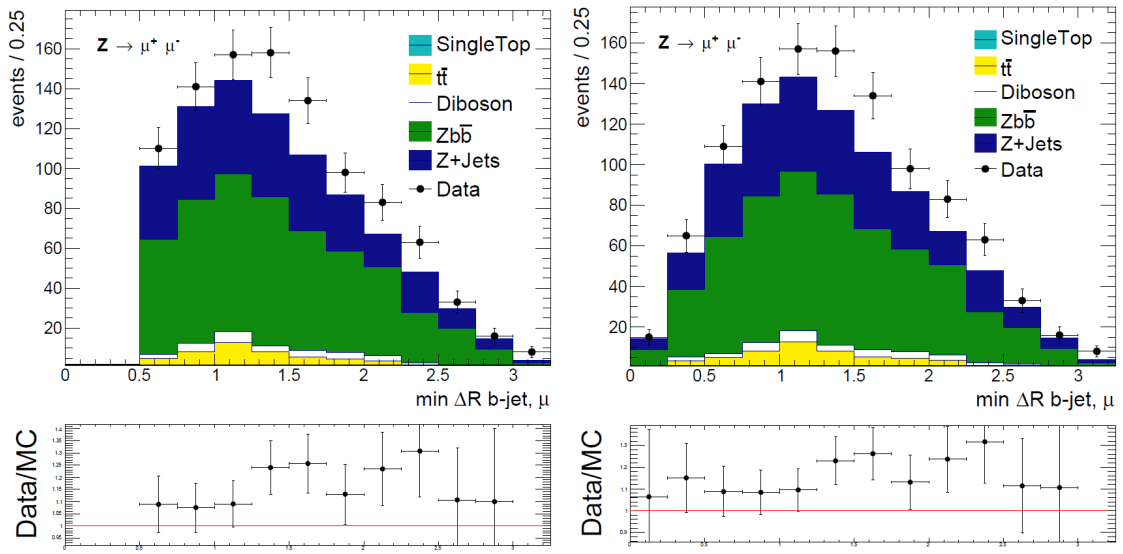


Figure B.11: Minimum signal muon–selected jet ΔR distributions with (left) and without (right) applying the Overlap Removal between muons and jets.

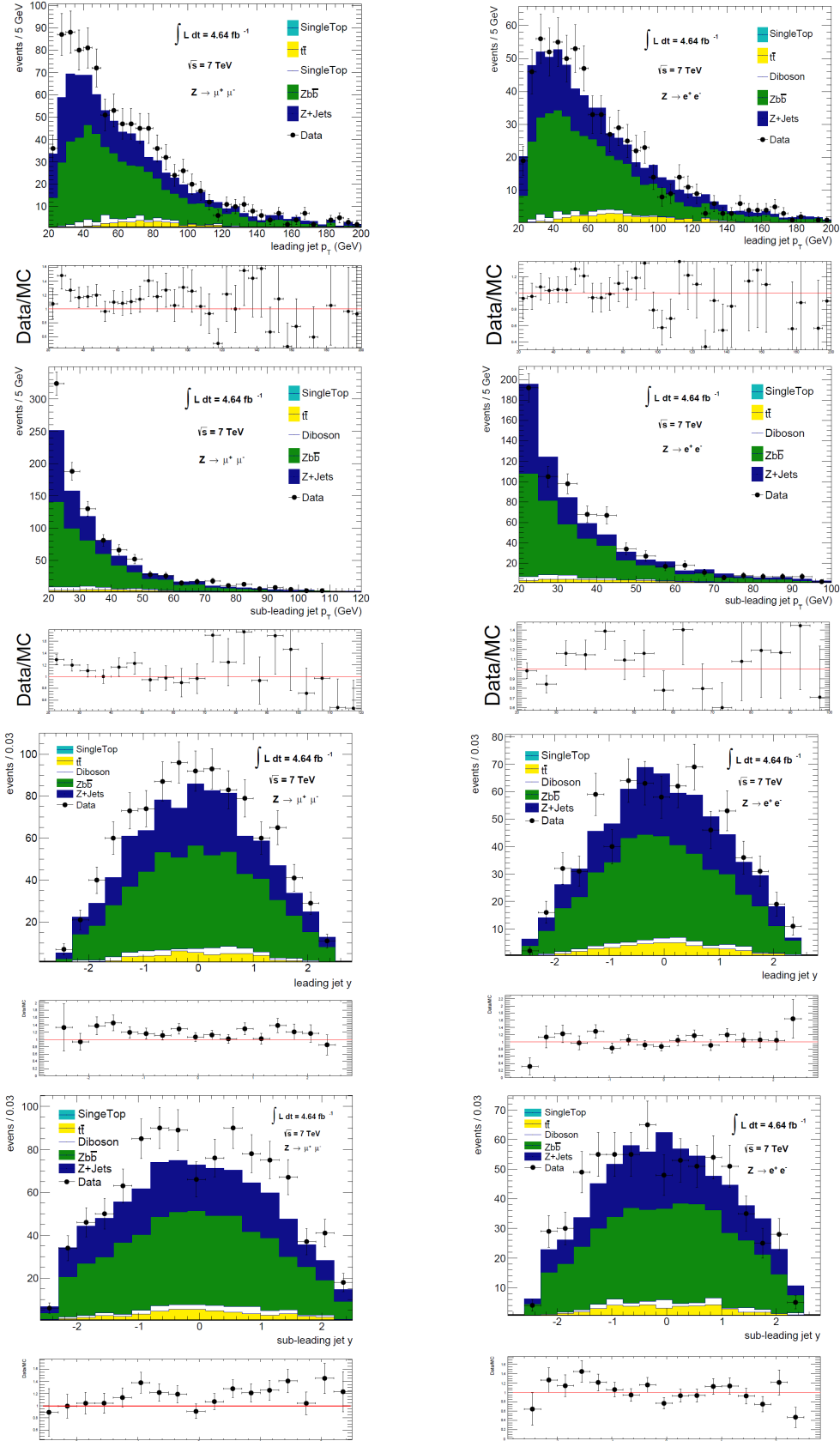


Figure B.12: Data and simulation distributions for the leading (first and third rows) and sub-leading (second and fourth rows) jets. The top (bottom) two rows show the transverse momentum (rapidity) distributions. Left (right) side distributions correspond to the electron (muon) channel.

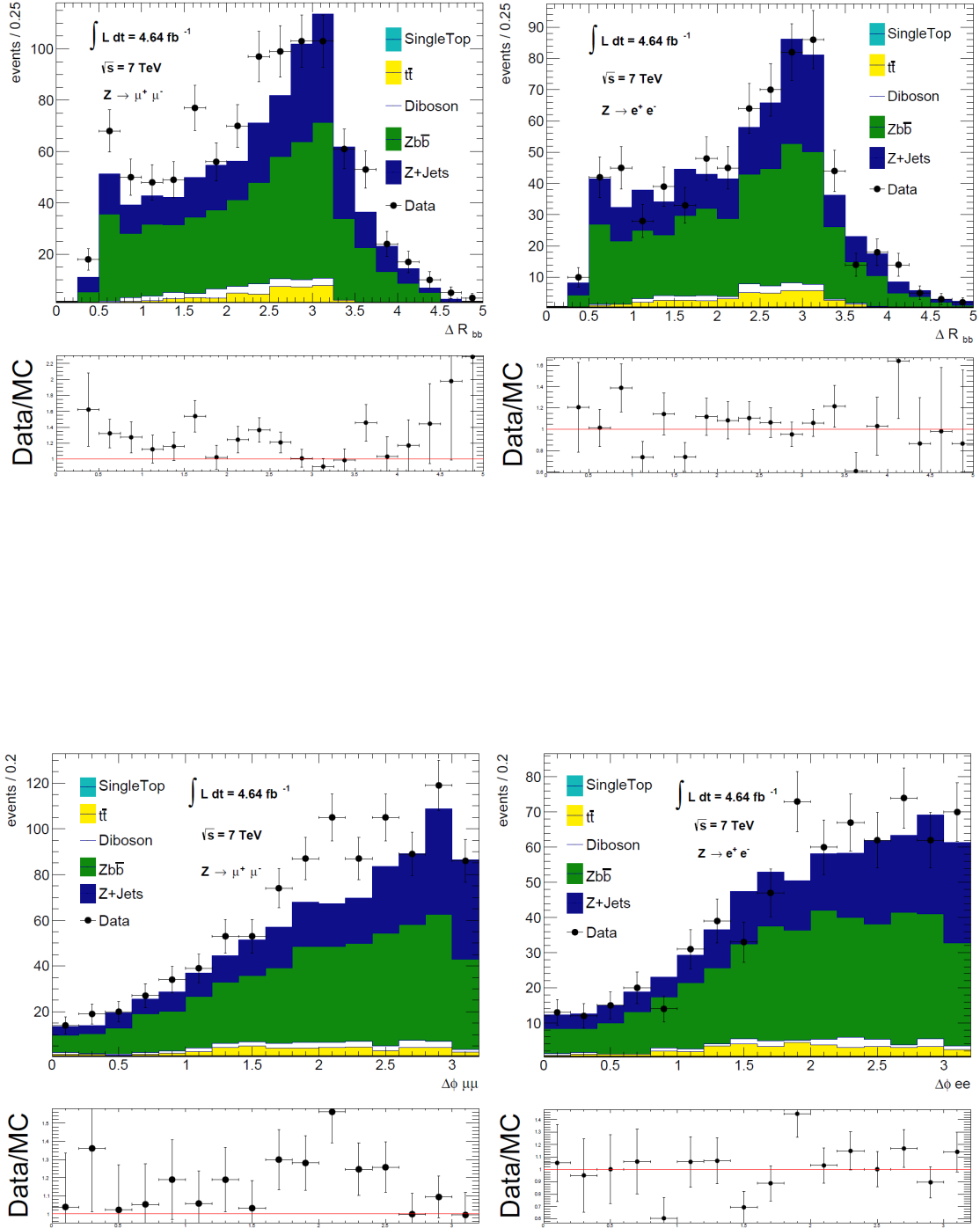


Figure B.13: Data and simulation ΔR ($\Delta\Phi$) distributions of the leading and sub-leading selected jets (leptons) are shown in the top (bottom) row. Left (right) side distributions correspond to the muon (electron) channel.

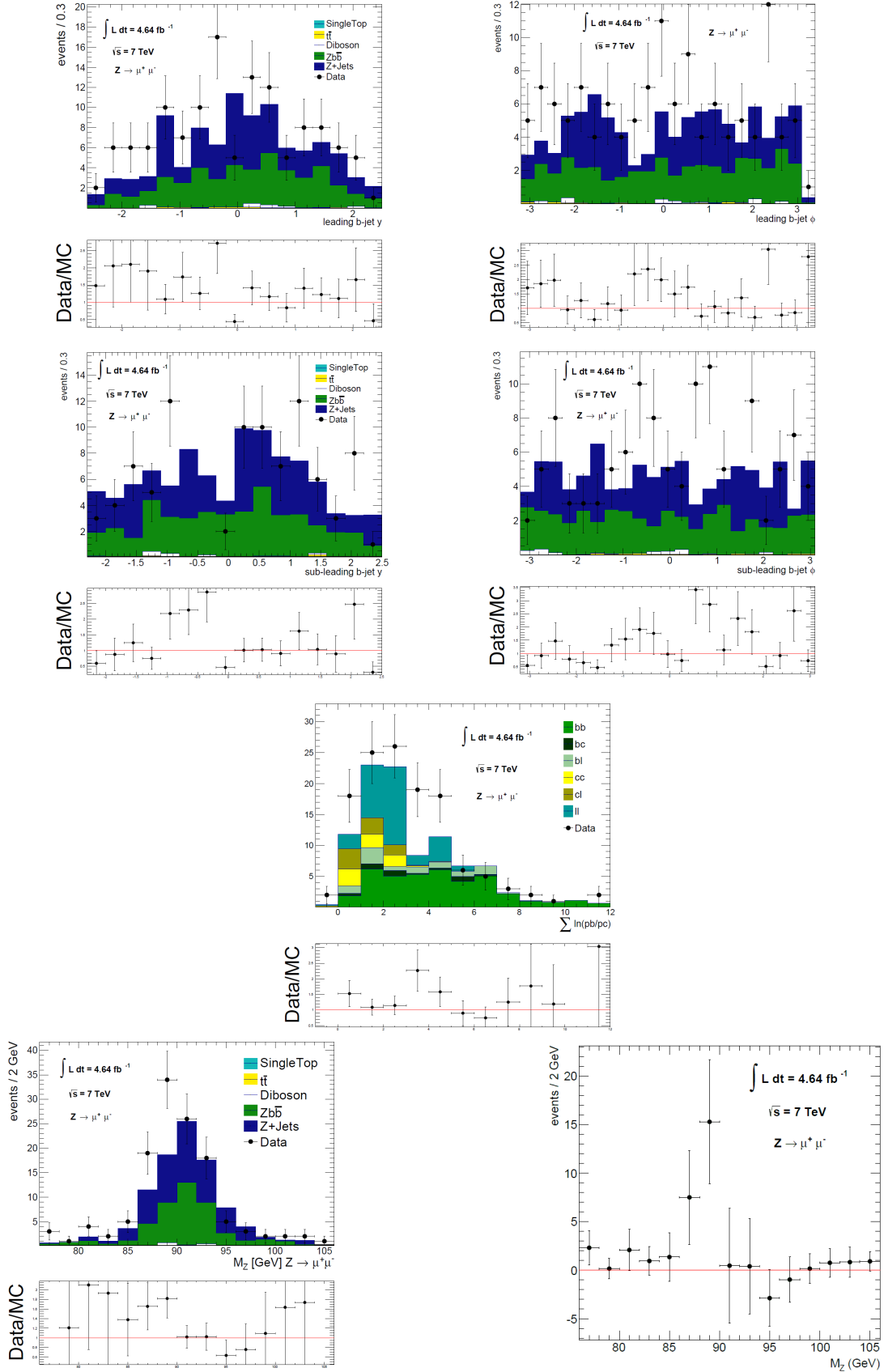


Figure B.14: Rapidity (first and second rows, left) and ϕ (first and second rows, right) angular distributions for the leading (first row) and sub-leading (second row) jets. Pre-fit template distribution splitted into individual jet flavor is shown on the third row. Signal muons invariant mass distributions for data and MC (bottom row, left) and data excess (bottom row, right).

Bibliography

- [1] Peter W. Higgs, Broken symmetries, massless particles and gauge fields, *Physics Letters* 12, 132 (1964)
- [2] Peter W. Higgs, Broken Symmetries and the masses of gauge bosons, *Phys. Rev. Lett.* 13, 508 (1964).
- [3] F. Englert, R. Brout, Broken Symmetry and the Mass of Gauge Vector Mesons, *Phys. Rev. Lett.* 13, 321 (1964).
- [4] The ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, *Phys. Lett. B* 716 (1): 1-29 (2012).
- [5] The CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, *Phys. Lett. B* 716 (1): 30-61 (2012).
- [6] The ATLAS Collaboration, Study of the spin of the new boson with up to in 25 fb^{-1} of ATLAS data, *ATLAS-CONF-2013-040*.
- [7] The ATLAS Collaboration, Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data, *ATLAS-CONF-2013-034*.
- [8] J. Beringer et al. (Particle Data Group,) *Phys. Rev. D* 86, 010001 (2012).
- [9] F. Halzen, A.D. Martin. Quarks & Leptons, John Wiley Sons, Inc., 1984.
- [10] David Griffiths. Introduction to Elementary Particles, WILEY-VCH Verlag GmbH
- [11] A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, Parton distributions for the LHC, *Eur.Phys.J.* C63:189-285,2009.
- [12] ATLAS Experiment Public Results. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#2011_pp_Collisions, URL last visited: Feb. 2013.

- [13] *LHC Higgs Cross Section Working Group* <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections>, URL last visited: Feb. 2013.
- [14] T. Sjöstrand, S. Mrenna, and P. Skands. *Pythia 6.4 physics and manual*. JHEP, 0605:026,2006.
- [15] G. Corcella et al. *Herwig 6.5: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes)*. JHEP, 2001, 010, 2001.
- [16] M.L. Mangano et al. *Alpgen, A generator for hard multiparton processes in hadronic collisions*. JHEP, 0307 (2003) 001.
- [17] S. Frixione and B.R. Webber. *The MC@NLO 3.2 event generator*, arXiv:hep-ph/0601192, 2006.
- [18] T. Gleisberg et al. *Event generation with Sherpa 1.1.*, JHEP 02 (2009)007.
- [19] Butterworth, J et al. *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7\text{TeV}$* ATL-COM-PHYS-2010-695.
- [20] *LHC Higgs Cross Section Working Group* <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt7TeV>, URL last visited: Feb. 2013.
- [21] *The ATLAS Experiment*, <http://public.web.cern.ch/public/en/research/AccelComplex-en.html>, ATLAS Experiment © 2012 CERN, URL last visited: Feb. 2013.
- [22] *The ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 S08003 (2008).
- [23] *The ATLAS Experiment* <http://www.atlas.ch/photos/calorimeters-combined-barrel.html>, ATLAS Experiment © 2012 CERN, URL last visited: Feb. 2013.
- [24] *The ATLAS Collaboration, Commissioning of the ATLAS Muon Spectrometer with Cosmic Rays*, Eur.Phys.J.C70:875-916,2010.
- [25] *The ATLAS Collaboration, The ATLAS Muon Spectrometer Technical Design Report*, CERNLHCC/97-22, 1997
- [26] *The ATLAS Experiment* <http://www.atlas.ch/photos/muons-combined.html>, ATLAS Experiment © 2012 CERN, URL last visited: Feb. 2013.

- [27] A. Yamamoto *et al.* *The ATLAS central solenoid*, Nuclear Instruments and Methods in Physics Research Section A 584 (1) 53-74 (2008).
- [28] *The ATLAS Collaboration*, *Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* (2011) 71: 1630
- [29] *Worldwide LHC Computing Grid*. <http://lcg.web.cern.ch/lcg/>, URL last visited: Feb. 2013.
- [30] *W/Z and Electroweak Common Topics (2011)*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/WZElectroweakCommonTopics2011>, URL last visited: Feb. 2013.
- [31] *ATLAS Internal, Heavy Flavor Overlap Removal Tool*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HforTool>, URL last visited: Feb. 2013.
- [32] *ATLAS Internal, Top group's MC11(a,b,c) Samples For 2011 Data Analyses*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC11#ttbar_Samples_in_MC11_Production, URL last visited: Feb. 2013.
- [33] J. Alwall *et al.* *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *Eur.Phys.J.C* 53:473-500,2008
- [34] *The ATLAS Collaboration*, *Reconstruction of primary vertices in pp collisions at energies of 900 GeV and 7 TeV with the ATLAS detector*, ATLAS-CONF-2010-069
- [35] *ATLAS Internal, ATLAS ElectronGamma Physics Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronReconstruction>, URL last visited: Feb. 2013.
- [36] *The ATLAS Collaboration*, *Observation of inclusive electrons in the ATLAS experiment at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-073
- [37] *The ATLAS Collaboration*, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011*, ATLAS-CONF-2012-048
- [38] *ATLAS Internal, ATLAS ElectronGamma Physics Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EfficiencyMeasurements>, URL last visited: Feb. 2013.
- [39] *ATLAS Internal, ATLAS ElectronGamma Physics Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronsEnergyDirection>, URL last visited: Feb. 2013.

- [40] *ATLAS Internal, ATLAS ElectronGamma Physics Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EnergyScaleResolutionRecommendations>, URL last visited: Feb. 2013.
- [41] *ATLAS Internal, ATLAS Muon Combined Performance*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MuonCombinedEventDataModel#AnchorMuonObject>, URL last visited: Mar. 2013.
- [42] *The ATLAS Collaboration, Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, ATLAS-CONF-2011-063
- [43] *ATLAS Internal, ATLAS Muon Combined Performance*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesRel17MC11a>, URL last visited: Mar. 2013.
- [44] *Matteo Cacciari et al. The anti- k_t jet clustering algorithm*, JHEP 0804:063,2008
- [45] *Lampl, W et al. Calorimeter Clustering Algorithms : Description and Performance*, ATL-LARG-PUB-2008-002
- [46] *ATLAS Internal, ATLAS Jet and Missing Et Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JESUncertaintyProvider?rev=54>, URL last visited: Mar. 2013.
- [47] *ATLAS Internal, ATLAS Jet and Missing Et Group*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets2011#Bad_jets, URL last visited: Mar. 2013.
- [48] *The ATLAS Collaboration, Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102
- [49] *The ATLAS Collaboration, Calibrating the b -Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector*, ATLAS-CONF-2011-089
- [50] *The ATLAS Collaboration, Measurement of the b -tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector*, ATLAS-CONF-2012-043
- [51] *The ATLAS Collaboration, Performance of the ATLAS Secondary Vertex b -tagging Algorithm in 7 TeV Collision Data*, ATLAS-CONF-2010-042
- [52] *ATLAS Internal, ATLAS Flavour Tagging Working Group*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTaggingCalibrationDataInterface>, URL last visited: Feb. 2013.

- [53] *The ATLAS Collaboration, Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS, Eur.Phys.J.C 72 (2012) 1844*
- [54] *The ATLAS Collaboration, Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7\text{TeV}$, Eur.Phys.J.C71:1577,2011*
- [55] *ATLAS Internal, ATLAS Standard Model Working Group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/WZCommonMETUtilityHowTo>, URL last visited: Feb. 2013.*
- [56] *ATLAS Internal, ATLAS Trigger and Data Acquisition, <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/TriggerObjectMatching>, URL last visited: Feb. 2013.*
- [57] *ATLAS Internal, W/Z and Electroweak Common Topics, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/WZElectroweakCommonTopics2011#MET_cleaning, URL last visited: Mar. 2013.*
- [58] *ATLAS Internal, ATLAS Flavour Tagging Working Group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/FlavourTagging>, URL last visited: Feb. 2013.*
- [59] *John M. Campbell et al., <http://mcfm.fnal.gov/>, URL last visited: Feb. 2013.*
- [60] *The ATLAS Collaboration, Improved Luminosity Determination in pp Collisions at $\sqrt{s} = 7\text{ TeV}$ using the ATLAS Detector at the LHC, ATLAS-CONF-2012-080*
- [61] *ATLAS Internal, ATLAS Jet and Missing Et Group, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetEnergyResolutionProvider>, URL last visited: Mar. 2013.*
- [62] *P. M. Nadolsky et al, Implications of CTEQ global analysis for collider observables, Phys.Rev.D78:013004, 2008*
- [63] *ATLAS Internal, ATLAS Standard Model Working Group, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/WZElectroweakCommonTopics2011#Reweight_vertex_z_position_on_MC, URL last visited: June. 2013.*
- [64] *D. Hellmich, Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universitaet Bonn, 2013*
- [65] *D.J. Lange Nucl.Instrum.Meth. A462 (2001) 152-155*
- [66] *The ATLAS Collaboration ATL-COM-PHYS-2012-1504*

Acknowledgements

I would like to begin by expressing my deepest gratitude to Prof. Dr. Dorothee Schaile for offering me a Ph.D. position in her group as well as for the financial support of the numerous trips to CERN, summer schools and conferences I've participated to over the past three years.

I also wish to thank Prof. Dr. Martin Faessler for his willingness to write the second review of my thesis. To Prof. Dr. Andreas Burkert, Prof. Dr. Hans Böhringer and Prof. Dr. Wolfgang Dünneweber I wish to thank for having agreed to participate in my Ph.D. defence committee.

For the numerous hours of discussions, constructive suggestions, the choice of my thesis subject, a door always open and for patiently correcting several draft versions of this thesis, I thank my direct coordinator Dr. Michiel Sanders, *Heel erg bedankt!*

A special thanks to Prof. Dr. Otmar Biebel, spokesman of the Graduiertenkolleg for the financial support of the summer schools I've participated to and for organising many interesting events in association with the International Max Planck Research School on Elementary Particle Physics.

I feel profoundly indebted to Prof. Dr. Ionel Lazanu, Prof. Dr. Alexandru Jipa, Prof. Dr. Mihaela Sin for writting my Ph.D. recommendation letters and especially Prof. Dr. Ionel Lazanu and Dr. Cristina Galea for their support.

For helping me become an ATLAS author I want to thank Dr. Stephanie Zimmermann.

To our analysis collaborators, Sonja Hillert, Dennis Hellmich and the ATLAS Z+b(b) groups I wish to thank for their many good ideas and suggestions.

To PD Dr. Johannes Elmsheuser I would like to thank for the intelligent comments and observations during my Abteilungseminar talks and for all the grid-related help.

Sincere thanks to our former secretary Frau Herta Franz and to our present secretary Frau Elke Grimm-Zeidler for all the paper-work help and good advices.

For the help with the writing of the German abstract of my thesis I wish to thank Dr. Julien de Graat.

These three and a half years would certainly not have been the same without my office colleagues Dr. Raphael Mameghani, Dr. Markus Lichtnecker and Dipl.-Phys. Christopher Schmitt to whom I owe lots of pleasant moments and enjoyable discussions.

For making my Ph.D. stage in Munich enjoyable I wish to thank my former and present Lehrstuhl colleagues.

To my close family members, I want to thank for their constant encouragement and support over the years.

Last but not least I want to thank my fiancée Magda for her endless patience during the writing of this thesis.

